



الجمهورية الجزائرية الديمقراطية الشعبية
RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE



Université des Frères Mentouri Constantine1
Faculté des Sciences de la Nature et de la Vie

جامعة الاخوة منتوري قسنطينة1
كلية علوم الطبيعة والحياة

Département : **Biologie Animale** قسم : **بيولوجيا الحيوان**

Mémoire présenté en vue de l'obtention du **Diplôme de Master**

Domaine : **Sciences de la Nature et de la Vie**

Filière : **Sciences Biologiques**

Spécialité : **Génétique**

Intitulé :

***Etude in silico des mutations de la protéine S « spike protein »
détectées dans les principaux variants du SARS-CoV-2 aperçus
en Algérie.***

Présenté et soutenu par : **KSIR NAZIM**

Le : **20 /09/2021**

Jury d'évaluation :

Président du jury : *SEDRATI KHADIDJA* (Maitre de conférences classe A – UFMConstantine 1).

Rapporteur : *OTMANI KHAWLA* - Maitre de recherche classe B – CRSP

Examineurs : *GUENDOUZE ASSIA* (Grade Maitre de conférences classe A – UFMConstantine 1).

***Année universitaire
2020- 2021***

Remerciements

Aux membres du jury, pour l'honneur qu'ils nous ont fait de participer à ce jury. Je tiens également à les remercier pour le temps qu'ils ont consacré à la lecture et la correction de ce modeste travail.

A Madame SEDRATI KHADIDJA qui nous a fait l'honneur d'accepter la présidence du jury.

Veillez trouver ici l'expression de mes sincères remerciements.

A Madame GUENDOUZE ASSIA qui nous a fait l'honneur de juger notre travail. Veuillez trouver

ici l'expression de mes sincères remerciements.

A Madame OTHMANI KHAWLA, qui m'a fait l'honneur d'accepter de m'encadrer, d'avoir corrigé et apporter une aide précieuse au cours de l'élaboration de ce travail. Je tiens tout particulièrement

à la remercier pour sa gentillesse, son implication, sa compétence, ses précieux conseils, sa disponibilité.

Aux personnels du centre de recherche en science pharmaceutique qui ont fait preuve d'une hospitalité et d'une gentillesse sans égale.

A Monsieur MOHAMED ABDELHAFID HAMIDECHI, d'avoir eu l'amabilité de prendre de son précieux temps pour m'expliquer un peu la bio-informatique, une discipline qui m'était complètement étrangère, je vous remercie infiniment et n'oublierai jamais votre gentillesse.

Enfin, merci à tous les professeurs de la spécialité GÉNÉTIQUE

Dédicaces

Je dédie ce modeste travail à mes parents qui m'ont soutenue et encouragé dans mes choix,
sans eux je n'aurais sans doute pas pensé à reprendre mes études. Je vous souhaite une
longue vie accompagnée d'une bonne santé.

À ma petite sœur à qui je dois l'idée d'inscription au Master 2 en génétique, je te souhaite
beaucoup de succès dans ta vie.

A tout le reste de la famille ; mon grand-père, mes deux grands frères (Djamil et Rouf), mes
oncles et tantes. Je vous souhaite prospérité et santé.

Sommaire

Remerciements

Dédicaces

Liste des abréviations

Table des illustrations

Résumé

Introduction 1

Partie bibliographique

1	Généralité sur les coronavirus	3
1.1	Définition des coronavirus	3
1.2	Taxonomie des coronavirus	4
1.3	Historique des coronavirus	7
2	Généralité sur la COVID19	8
2.1	SARS-COV-2	8
2.1.1	Définition de la COVID19 « SARS-COV-2 »	8
2.1.2	Historique de la COVID19	9
2.1.3	Epidémiologie du SARS-COV-2	10
2.1.3.1	Distribution géographique du SARS-COV-2	10
2.1.4	Origine du SARS-COV-2	14
2.2	Les variants du SARS-COV-2	15
2.2.1	Définitions des variants du SARS-COV-2	15
2.2.2	Nomenclature des variants du SARS-COV-2	16
2.2.3	Historique des variants du SARS-COV-2	17
2.2.4	Epidémiologie des variants du SARS-COV-2	18
2.2.4.1	Dans le monde	18
2.2.4.1.1	Le variant Anglais « Alpha »	18
2.2.4.1.2	Le variant Sud-Africain « bêta »	19
2.2.4.1.3	Le variant Brésilien « Gamma »	19
2.2.4.1.4	Le variant Indien « Delta »	20
2.2.4.1.5	Le variant Nigérien « Eta »	21
2.2.4.2	En Algérie	21
2.3	Contagiosité du SARS-COV-2 et des variants	22
2.4	Taux de létalité du SARS-COV-2 et des variants	24
2.4.1	Virulence et variants du SARS-COV-2	24
2.4.2	Sexe	25
2.4.3	Age	26

3	Biologie du SARS-COV-2	27
3.1	Structure et organisation génomique des coronavirus	27
3.2	Organisation génomique du SARS-COV-2	28
3.3	Structure des protéines du SARS-COV-2	30
3.3.1	Protéine Spike « S »	30
3.3.2	La protéine nucléocapside « N »	32
3.3.3	La protéine membrane « M »	32
3.3.4	La protéine enveloppe « E »	33
3.3.5	Les protéines non structurales	33
3.4	Cycle de vie du SARS-COV-2	34
3.4.1	L'attachement et l'entrée dans la cellule	35
3.4.2	Traduction de l'ARN viral	35
3.4.3	Transcription et réplication génomique « RNA synthesis »	36
3.4.4	Traduction des protéines structurales, assemblage et libération des virions	38
3.5	Les variants du SARS-COV-2	38
3.5.1	Le variant G614	38
3.5.2	Les variants préoccupants « variant of concern »	39
3.5.2.1	Le variant Anglais « Alpha »	39
3.5.2.2	Le variant Sud-Africain « Béta » :	41
3.5.2.3	Variant Brésilien « Gamma »	42
3.5.2.4	Le variant Indien « Delta »	43
3.5.3	Les variants à suivre (variant of interest)	43
3.5.3.1	Les variants Américains	43
3.5.3.1.1	Le variant Epsilon	43
3.5.3.1.2	Le variant Iota « B.1.526 »	44
3.5.3.2	Le variant Nigérien « Eta »	44
3.5.4	Les variants en cours d'évaluation (variant under monitoring)	44
3.5.4.1	Le variant B.1.214.2 :	45
3.5.4.2	Le variant N.9 :	45
3.6	Pathogénie du SARS-COV-2	46
3.6.1	Le système rénine-angiotensine	46
3.6.2	Système rénine angiotensine et SARS-COV-2	46
3.6.3	Mécanisme d'échappement du SARS-COV-2 à l'immunité innée de l'hôte	47
3.7	Symptômes du SARS-COV-2	48
3.8	Physiopathologie du SARS-COV-2	50
3.9	Diagnostic	51
3.10	Traitement du SARS-COV-2	53
3.10.1	Les traitements palliatifs (formes modérées)	53
3.10.2	L'élimination virale à l'aide de molécules capables d'interférer avec la réplication virale	54
3.10.3	Les thérapies anti-inflammatoires stéroïdiens (formes graves)	54
3.10.4	L'azithromycine (traitement le plus utilisé en Algérie)	54

3.10.5	Les anti-hémostatiques	54
3.11	Vaccin antiviral	55
3.12	Impact des variants sur la pathogénie du SARS-COV-2	56
4	La bio-informatique et la COVID19	58
4.1	Généralité sur la bio-informatique	58
4.1.1	Bases de données	58
4.1.2	Banques de Données Biologiques	58
4.1.2.1	Les Banques de Séquences Nucléiques	59
4.1.2.1.1	EMBL	59
4.1.2.1.2	GenBank	59
4.1.2.1.3	DDBJ (Dna Data Bank)	59
4.1.2.2	Les Banques de Séquences Protéiques	59
4.1.2.2.1	PIR-NBRF	59
4.1.2.2.2	SWISSPROT	59
4.1.3	LES GENOME BROWSERS	60
4.2	Rôle de la bio-informatique dans l'étude de la physiopathologie du SARS COV2 et variants	60
4.3	Alignement des séquences	60
4.3.1	Alignement de deux séquences	61
4.3.2	L'Alignement Multiple des Séquences	61
5	Matériels et méthodes	64
5.1	Acquisition des séquences nucléotidiques et protéiques	64
5.1.1.1	Rechercher sur NCBI la séquence d'intérêt	64
5.1.1.2	Identifier la séquence souhaitée	65
5.1.1.3	Choisir le résultat de recherche qui correspond à notre recherche	66
5.1.1.4	Faire apparaître la séquence d'intérêt	67
5.1.1.5	Utiliser le lien FASTA pour faire apparaître la séquence protéique	68
5.1.1.6	Apparition de la séquence de la protéine S	69
5.1.1.7	Obtention des séquences de la protéine S des variants	69
5.2	L'Alignement Multiple des Séquences protéiques	70
6	Résultats et discussions	74
6.1	Résultats de l'alignement des séquences	74
6.1.1	Délétion H69/V70	77
6.1.2	Mutation K417N/T	78
6.1.3	Mutation L452R	79
6.1.4	Mutation 478K, E484K/Q et N501Y	80
6.1.5	La mutation D614G	83
6.2	Pourcentage d'identité	84
7	Conclusion	87

Références bibliographiques

Liste des abréviations

229E : Specimen code 229E

ACE2 : Angiotensin Converting Enzyme 2

AINS : Anti Inflammatoire Non Stéroïdiens

APN: Amino Peptidase N

AT1 (AT1R): Angiotensin II Receptor type 1

BCoV: Bat CoronaVirus

CCoV: Coronavirus Canin

CEACAM1: Carcinoembryonic Antigen Cell Adhesion Molecule 1

DDBJ: DNA Data Bank Japan

EBI: European Bioinformatics Institute

Elisa: Enzyme-Linked Immunosorbent Assay

EMBL: European Molecular Biology Laboratory

FeCoV: Feline CoronaVirus

FIPV: Feline Infectious Peritonitis Virus

FP: Fusion Peptide

GISAID: Global Initiative on Sharing Avian Influenza Data.

HCoV: Human Corona Virus

HKU1: Hong Kong University 1

HR: Heptad Repeat

IBV: Avian coronavirus

JIPID : Japan International Protein Information Database

LAMP: Loop-Mediated Isothermal Amplification

MAS: Mas-related G protein-coupled receptors

MDA-5: Melanoma Differentiation-Associated Protein 5

MERS: Middle EAST Respiratory Syndrome.

MHV: *Murine Hepatitis Virus*)

MIPS: Martinsried Institute for Protein Sequences

Muscle: Multiple Sequence Comparison by Log-Expectation

NBRF: National Biomedical Research Foundation

NCBI: National Center for Biotechnology Information

Neu5,9Ac2: Acide N-acétylneuraminique

NF-KB: Nuclear Factor Kappa-light-chain-enhancer of activated B cells

NL63 : Netherland 63

NTD : N-Terminal Domain

OC43 : Organ Culture 43

OMS : Organisation Mondiale de la Santé

ORF: Open Read Frame

PEDV: Porcine epidemic diarrhea virus

PLP: Papain-Like Protease.

PNS: Protéines Non Structurales

POCT: Point of Care Tests.

PP1a: Poly-Protein 1a

PRR: Pattern Recognition Receptor

RATG13: Bat *Rhinolophus affinis*, Tongguan, 2013

RBD: Receptor Binding Domain.

RBM: Receptor Binding Motif

RdRp: RNA Dependent RNA Polymerase

RIG-1: Retinoic Acid-inducible Gene I

RT-PCR: Reverse Transcriptase PCR- Polymerase Chain Reaction

SARS-COV-2: Severe Acute Respiratory Syndrome.

SGTF: S Gene Target Failure

SRA: Système Rénine Angiotensine

STAT1 Signal Transducers and Activators of Transcription

TDM: Tomodensitométre

TGEV: Transmissible Gastroenteritis Virus

TMPRSS-2: Transmembrane Protease Serine 2

TM: Transmembran Motif

TRIM25: Tripartite Motif Containing 25

UCSC: University of California Santa Cruz

UTR : Untranslated Regions

Table des illustrations

Figures

Partie bibliographique

Figure 1 : Distribution mondiale des coronavirus humains.

Figure 2 : Taxonomie des coronavirus.

Figure 3 : Morphologie des coronavirus au microscope électronique.

Figure 4 : Aspect structurelle des coronavirus.

Figure 5 : distribution du SARS-COV-2 dans le monde.

Figure 6 ; évolution de la pandémie dans la province de Anhui, Chine.

Figure 7 : Cas confirmés du SARS-COV-2 déclarés en Afrique au 08 septembre 2021.

Figure 8 : Évolution quotidienne des cas de SARS-COV-2 déclarés en Algérie entre le 15 février 2020 et le 07 septembre 2021.

Figure 9 : les hôtes naturels et intermédiaires des Coronavirus.

Figure 10 : Distribution géographique du variant Anglais.

Figure 11 : Distribution géographique du variant Sud-Africain.

Figure 12 : Distribution géographique du variant Brésilien.

Figure 13 : Distribution géographique du variant Indien.

Figure 14 : Les variants détectés en Afrique et en Algérie au 27 aout 2021.

Figure 15 : Schéma expliquant le taux de reproduction de base R_0 .

Figure 16 : Comparaison du R_0 des variants avec la souche originelle du SARS-COV-2.

Figure 17 : Taux de létalité selon l'âge, le sexe et les conditions de santé.

Figure 18 : variation de la réponse immunitaire selon le sexe.

Figure 19 ; Structure et organisation du génome des coronavirus.

Figure 20 : comparaison de l'organisation génomiques des coronavirus.

Figure 21 : Organisation du génome du SARS-CoV2.

Figure 22 : pourcentage de similarité du génome du SARS-COV-2 avec les autres coronavirus.

Figure 23 : Schéma illustrant la liaison entre le SARS-CoV-2 et le récepteur ACE2 présent à la surface des cellules cibles du coronavirus (cellule de l'hôte).

Figure 24 : structure de la protéine S du SARS-COV-2.

Figure 25 : Structure en trimère de la protéine S.

Figure 26 : similitude de la structure génomique de la protéine S des coronavirus.

Figure 27 : structure des protéines E, M, N.

Figure 28 : les protéines non structurales des coronavirus.

Figure 29 : Cycle de vie et répllication virale du SARS-COV-2 dans la cellule hôte.

Figure 30 : Mécanisme de l'entrée du SARS-COV-2 à l'intérieur de la cellule hôte.

Figure 31 : organisation génomique et transcription virale du SARS-COV-2.

Figure 32 : schéma détaillant les différentes étapes de synthèse de l'ARN viral.

Figure 33 : Structure du complexe RdRp. « ARN polymérase ARN dépendante ».

Figure 34 : La mutation D614G au niveau de la protéine S du SARS-COV-2.

Figure 35 : les mutations clés présentes dans le génome du variant Anglais.

Figure 36 : les mutations de la protéine S du variant Anglais.

Figure 37 : les mutations observées dans les variants « Alpha, Béta, Gamma ».

Figure 38 : Les principales mutations observées dans le SARS-COV-2.

Figure 39 : Rôle du SRA dans la virulence du SARS-COV-2.

Figure 40 : mécanisme de la détection immunitaire et de l'évasion virale.

Figure 41 : les manifestations cliniques du SARS-COV-2.

Figure 42 : Complications après l'infection au SARS-COV-2.

Figure 43 : physiopathologie du SARS-COV-2

Figure 44 : les différentes méthodes de diagnostic du SARS-COV-2.

Figure 45 : alignement de deux séquences protéiques.

Partie pratique

Figure 1 : Moteur de recherche du site NCBI.

Figure 2 : filtrage des résultats de recherche

Figure 3 : sélection du résultat de recherche voulu.

Figure 4 : page du génome complet du SARS-COV-2 sur le site NCBI.

Figure 5 : identification NCBI de la séquence protéique de la protéine S.

Figure 6 : choix du format FASTA

Figure 7 : apparition de la séquence au format FASTA

Figure 8 : importer les séquences dans le logiciel JALVIEW.

Figure 9 : Séquences copiées dans la console du logiciel.

Figure 10 : séquences non encore alignées par l'algorithme d'alignement.

Figure 11 : résultat de l'alignement multiple des séquences protéiques.

Figure 12 : résultat de l'alignement multiple des séquences protéiques.

Figure 13 : longueur des séquences protéiques dans le logiciel JALVIEW.

Figure 14 : Délétion H69/V70 détectée dans les variants Alpha et Eta.

Figure 15 : Mutation K417N présente dans le variant Béta et K417T dans le variant Gamma.

Figure 16 : Mutation L452R détectée dans le variant Delta.

Figure 17 : Les mutations T478K (Delta), E484K (Béta, Gamma, Eta), E484Q (Delta) et N501Y (Alpha, Béta, Gamma).

Figure 18 : Séquence protéique du domaine RBD.

Figure 19 : Mutation D614G détectée dans les variants : Alpha, Béta Gamma, Delta, Eta et. G614

Figure 20 : alignement de séquence deux à deux et pourcentage d'identité.

Tableaux

Partie Bibliographique

Tableau 1 : : Classification des espèces des coronavirus.

Tableau 2 : Classification de la famille des coronavirus selon l'hôtes, le virus en cause, les pathologies provoquées et le types de réponses cellulaires.

Tableau 3 : épidémiologie du SARS-COV-2 selon les continents et leurs pays les plus touchés.

Tableau 4 : Hôtes naturels et intermédiaires des coronavirus infectant l'homme.

Tableau 5 : Nomenclature scientifique des variants du SARS-COV-2.

Tableau 6 : Nomenclature des variants du SARS-COV-2 selon l'OMS.

Tableau 7 : comparaison du taux de reproduction et létalité du SARS-COV-2 avec les virus des précédentes pandémies.

Tableau 8 : avantages et inconvénients des méthodes de détection moléculaires du Sars-CoV2.

Tableau 9 : les 7 vaccins dont l'utilisation est approuvée par l'OMS.

Tableau 10 : Caractéristiques Cliniques et virologiques des principaux variants du SARS-COV-2.

Partie pratique

Tableau 1 : numéro d'accession des séquences de la protéine S.

Tableau 2 : comparaison de la longueur des séquences protéiques de la protéine S.

Tableau 3 : pourcentage d'identité des différents domaines du SARS-COV-2 des variants par rapport à la souche ancestrale.

Résumé

L'émergence de la maladie du syndrome respiratoire aigu sévère à coronavirus 2 (SARS-COV-2) en Chine fin 2019 a provoqué une vaste épidémie mondiale et constitue un problème majeur pour la santé publique. Malgré la lenteur de l'évolution du SARS-COV-2 par rapport aux autres virus à ARN, sa transmission massive et rapide pendant la pandémie mondiale lui a permis d'évoluer par l'acquisition d'une diversité génétique importante qui a conduit à l'émergence de nombreux variants, parmi lesquelles certains possèdent des mutations inquiétantes qui semblent avoir un impact sur la transmissibilité du virus, les manifestations cliniques, la mortalité ainsi que sur l'évasion immunitaire qui constitue le principale frein au développement d'un vaccin efficace contre les différentes souches du SARS-COV-2. Actuellement le **GISAID** (Global Initiative on Sharing Avian Influenza Data) a répertorié plus de 250 mutations au niveau du génome du SARS-COV-2 dont 60 mutations au niveau de la protéine S, cette fréquence élevée de mutations au niveau de la protéine S reflète son importance d'où l'intérêt de l'étudier.

Le présent travail a pour objectif d'étudier la protéine S et les mutations détectées dans sa séquence génomique, pour cela nous allons utiliser les séquences protéiques de la protéine S des différentes souches du SARS-COV-2 présentes dans les bases de données génomiques et les traiter par un algorithme d'alignement multiple pour les aligner et les comparer entre elles et par rapport à la souche ancestrale. Les résultats sont assez explicites en eux même et sont conformes aux données scientifiques des recherches déjà effectuées à savoir que la fréquence de mutation de la protéine S est plus élevée par rapport aux autres domaines du génome du SARS-COV-2 et que les mutations qui s'y produisent tout particulièrement au niveau du domaine **RBD** sont les plus dangereuses et qui ont eu le plus d'impacts sur la transmissibilité, l'évasion immunitaire et sur l'efficacité des vaccins.

Mots clés :

Coronavirus, COVID19, SARS-COV-2, Variant, ACE2, Protéine S, RBD, alignement séquence.

Abstract

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) disease in China in late 2019 caused a large global epidemic and is a major public health concern. Despite the the SARS-COV-2 have a slow rate of evolution compared to other RNA viruses, its massive and rapid transmission during the global pandemic has allowed it to evolve by acquiring significant genetic diversity that has led to the emergence of numerous variants, some of which possess worrisome mutations that appear to have an impact on virus transmissibility, clinical manifestations, mortality, as well as immune evasion, which is the main obstacle to the development of an effective vaccine against the different strains of SARS-COV-2. Currently, the GISAID (Global Initiative on Sharing Avian Influenza Data) has listed more than 250 mutations in the SARS-COV-2 genome, including 60 mutations in the Spike protein. This high frequency of mutations in the S protein reflects its importance, which is why it is worth studying it.

The objective of the present work is to study the S protein and the mutations detected in its genomic sequence, for that we will use the protein sequences of the S protein of the different strains of SARS-COV-2 present in the genomic databases and process them by a multiple alignment algorithm to align and compare them among themselves and with the ancestral strain. The results are quite self-explanatory and are consistent with scientific data from previous research which is that the mutation frequency of protein S is higher than other domains of the SARS-COV-2 genome and that the mutations that occurs in the RBD domain in particular are the most dangerous and have the greatest impact on transmissibility, immune evasion and vaccine efficacy.

Key words:

Coronaviruses, COVID19, SARS-COV-2, Variant, ACE2, Protein S, RBD, sequence alignment.

ملخص

انطلاقاً من منطقة اوهان في الصين، وذلك في أواخر عام 2019 تسبب فيروس كورونا SARS-COV-2 في انتشار وباء عالمي وبظهور مشكل صحي عويص ” مرض المتلازمة التنفسية الحادة”.

ورغم التطور البطيء لـ SARS-COV-2 مقارنة بفيروسات الحمض النووي الأخرى، إلا أن انتقاله كان بشكل سريع خلال فترة هذا الوباء، وذلك ما سمح له بالتطور من طريق اكتساب تنوع جيني مختلف و متعدد أدى إلى ظهور العديد من المتغيرات و الطفرات ، وبعض هذه المتغيرت له تأثير سلبي على قابلية انتقال الفيروس ، المظاهر السريرية ، والوفيات وكذلك على التحايل على الجهاز المناعي الذي يشكل عقبة رئيسة أمام تطوير لقاح فعال ضد سلالات مختلفة من SARS-COV-2

حالياً ، أدرجت GISAID (المبادرة العالمية لمشاركة بيانات إنفلونزا الطيور) أكثر من 250 طفرة في جينوم SARS-COV-2 ، بما في ذلك 60 طفرة في بروتين S ، ويعكس هذا التردد العالي للطفرات في بروتين S أهميته و قوة تأثيره (خطورته، او قدرته) وذلك مازاد من توجه الدراسات صوبه.

الهدف من هذا العمل هو دراسة بروتين S وكذا الطفرات المكتشفة في تسلسله الجينومي ، ولهذا سوف نستخدم تسلسل البروتين لبروتين S لسلالات مختلفة من SARS-COV-2 الموجودة في قواعد البيانات الجينومية ومعالجتها. من خلال خوارزمية محاذاة متعددة ومقارنتها مع بعضها البعض. خلصنا إلى جملة من النتائج الواضحة والتي تتوافق مع البيانات العلمية من البحوث الذي تم إجراؤها ، أي أن تواتر طفرة البروتين S أعلى مقارنة بالمناطق الأخرى من جينوم SARS-COV-2 وأن الطفرات التي تحدث بشكل خاص في مجال RBD هي الأكثر خطورة وكان لها أكبر تأثير على قابلية الانتقال والتحايل على الجهاز المناعي وفعالية اللقاح.

الكلمات الدالة:

فيروسات كورونا، سارس كوف 2، السلالة، الإنزيم المحول للأنجيوتنسين 2، بروتين إس، تراصف تسلسلي، نطاق الارتباط

Introduction

La COVID19 est une nouvelle maladie virale hautement contagieuse due à un virus à ARN le « SARS-CoV-2 » appartenant à la famille des coronavirus et qui est responsable de la troisième pandémie mondiale à coronavirus au cours de ces vingt dernières années (1). Les premiers cas de cette maladie virale à prédominance respiratoire ont été signalés à Wuhan dans la province chinoise du Hubei vers fin décembre 2019, puis le virus s'est rapidement propagé dans le monde entier et a ravagé de nombreux pays et a submergé de nombreux systèmes de santé obligeant l'Organisation mondiale de la santé (OMS) à déclarer la pandémie mondiale le 11 mars 2020(2). La COVID19 a eu un effet catastrophique sur la démographie mondiale entraînant plus de 4,5 millions de décès et plus de 220 millions de cas dans le monde en moins de deux ans, elle continue encore ses ravages à travers le monde et de nombreux pays font face à une deuxième ou troisième ou même une quatrième vague d'épidémie qui sont principalement attribuées à l'émergence de nouveaux variants du virus (3).

Comme d'autres virus à ARN, le SARS-CoV-2 est sujet à une évolution génétique avec le développement de mutations qui entraînent l'apparition de variants pouvant avoir des caractéristiques différentes de celles de ses souches ancestrales. Ainsi les variants peuvent être plus transmissibles, plus résistants et peuvent échapper à l'immunité en développant des mécanismes de contournement de la réponse immunitaire (4).

Dans le cadre de notre modeste étude, nous portons notre attention sur les mutations qui se sont produites au niveau de la protéine S car elle joue un rôle clé dans la pénétration du virus au sein des cellules de l'hôte, dans l'évasion immunitaire ainsi que sur la transmissibilité du virus. La protéine S est d'autant plus importante car elle constitue l'élément essentiel dans la fabrication des vaccins pour lutter contre le SARS-COV-2 (5).

L'objectif général de ce travail vise à comprendre les principales caractéristiques virales et cliniques du SARS-COV-2, à étudier les variants et comparer les séquences protéiques à l'aide d'un logiciel de multi alignement. Nous essaierons également de décrire l'impact des mutations sur la pathogénicité du virus, sur sa transmissibilité ainsi que sur l'efficacité des vaccins.

Partie
bibliographique

1 Généralité sur les coronavirus

1.1 Définition des coronavirus

Les coronavirus sont des agents infectieux courants qui infectent différents hôtes, y compris l'homme, ils provoquent des maladies respiratoires, gastro-intestinales, hépatiques et neurologiques. Ils sont une famille de virus à ARN enveloppés qui possèdent le plus grand génome de tous les virus à ARN décrits à ce jour (6).

Le nom des **coronavirus** qui signifie « **virus en couronne** » est dû à la présence de la glycoprotéine Spike sur l'enveloppe leur conférant ainsi l'aspect en couronne(7). Actuellement il existe 39 espèces qui appartiennent à la famille des Coronaviridae, parmi eux 7 coronavirus sont connus pour causer des infections respiratoires chez l'homme dont quatre pouvant provoquer des symptômes semblables à ceux du rhume : « **HCOV (Human Corona Virus) 229E (specimen code 229E), OC43 (Organ Culture 43), NL63 (Netherland 63), HKU1 (Hong Kong University 1)** ». Ceux qui infectent les animaux peuvent évoluer et devenir infectieux pour l'homme ainsi durant ces deux dernières décennies trois virus ont émergé : le SARS-COV-1 (Severe Acute Respiratory syndrome), le MERS (Middle EAST Respiratory syndrome) et le SARS-COV-2 (Figure 1)(6).

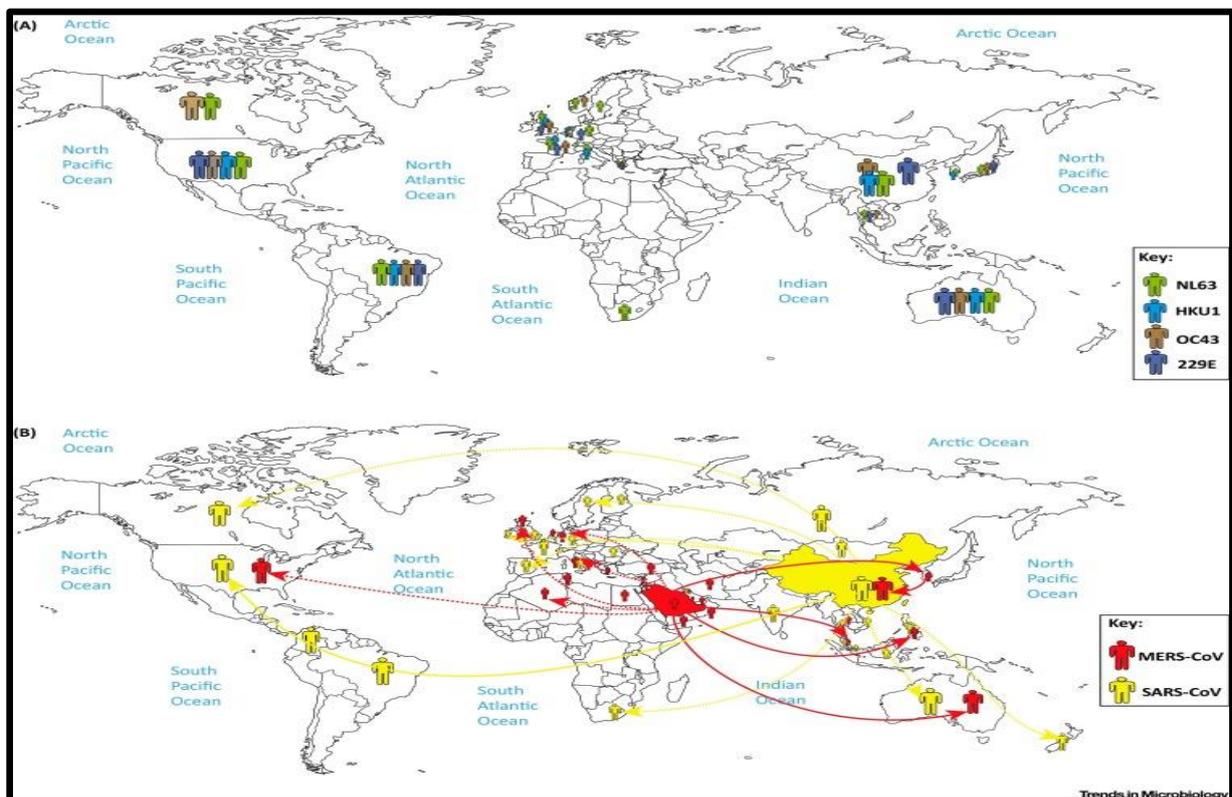


Figure 1 : Distribution mondiale des coronavirus humains (8)

1.2 Taxonomie des coronavirus

En 1968 Tyrrel a inventé le nom coronavirus qui est en référence à la forme caractéristique des coronavirus lorsqu'ils sont observés au microscope. Les coronavirus appartiennent à la sous famille des orthocoronavirinae, famille coronaviridae, de l'ordre nidovirale (9).

L'ordre des Nidovirales comprend trois familles ; les *Coronaviridae* (SARS-Cov-1 et 2, MERS), les *Arteriviridae* (pathogènes porcins et équins) et les *Roniviridae* (virus invertébrés). (Figure 2) (10).

Les virus appartenant à l'ordre Nidovirale possèdent des caractéristiques spécifiques communes (9) :

- L'expression génique par la transcription d'un ensemble de plusieurs segments ARN sous-génomiques imbriqués du côté de l'extrémité 3' formant un nid d'où le nom de l'ordre qui dérive du latin « nido » (9).
- L'expression des polyprotéines « **PP1a** et **PP1ab** » par décalage du cadre de lecture ribosomal (11).
- Possèdent un génome hautement conservé avec les gènes de réplication ce qui aide à la formation de gènes structurels et accessoires et de plusieurs activités enzymatiques uniques (12).
- Les virions sont enveloppés et contiennent une protéines membranaire intégrale « protéine M » (9).

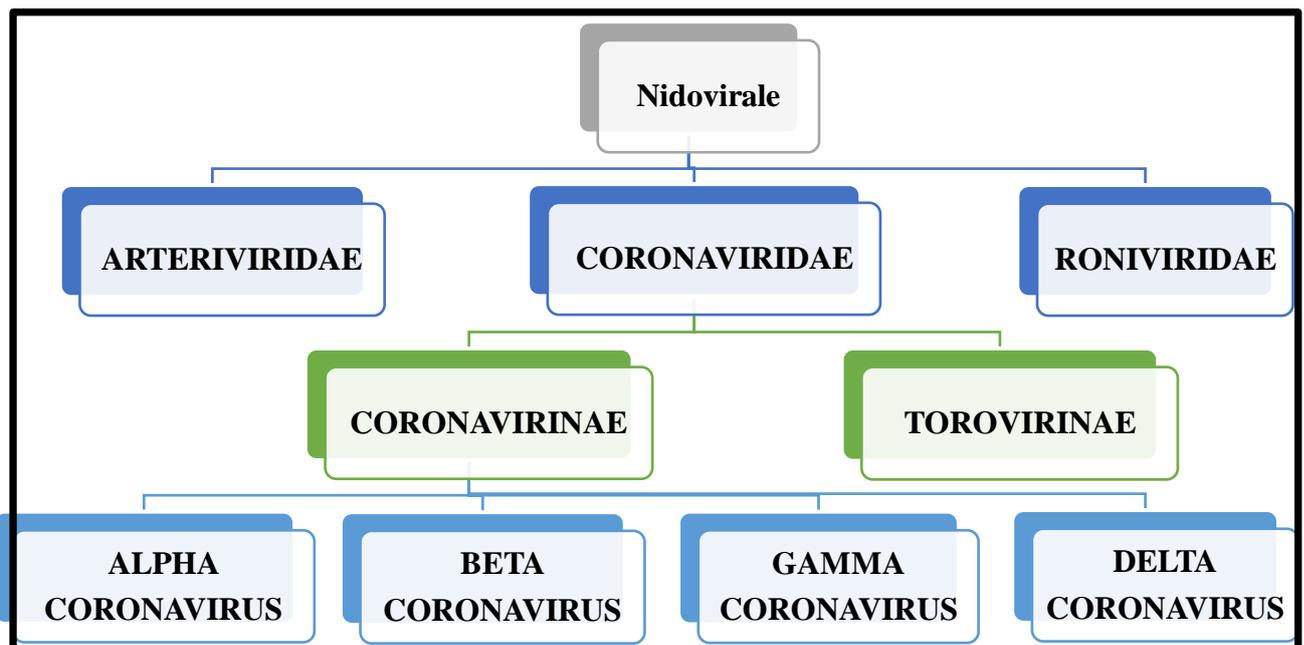


Figure 2 : Taxonomie des coronavirus (13).

La classification en groupes de la famille des coronaviridae était à l'origine basé sur les caractéristiques de leurs séquences protéiques et étaient donc classés en quatre genre (Tableau 1) :

- Alphacoronavirus « α -CoV »
- Bétacoronavirus « β -CoV »
- Gammacoronavirus « γ -CoV »
- Deltacoronavirus (14).

Tableau 1 : Classification des espèces des coronavirus (14).

Groupe	Espèces
α-Coronavirus	Coronavirus de la gastro-entérite transmissible (TGEV)
	Coronavirus Canin (CCoV)
	Coronavirus respiratoire porcin HKU15 (PRCoV)
	Coronavirus félin (FeCoV)
	Coronavirus de la diarrhée épidémique porcine (PEDV)
	Coronavirus humain 229E (HCoV-229E)
	Coronavirus humain NL63 (HCoV-NL63)
β- Coronavirus	Coronavirus de la chauves-souris (BCoV)
	Coronavirus de l'encéphalomyélite hémagglutinante porcine (HEV)
	Coronavirus Murin ou virus de l'hépatite murine (MHV)
	Coronavirus humain 4408 (HCoV-4408)
	Coronavirus humain OC43 (HCoV-OC43)
	Coronavirus humain HKU1 (HCoV-HKU1)
	Coronavirus du Syndrome respiratoire aigu sévère (SARS-COV)
	Coronavirus du syndrome respiratoire du Moyen-Orient (MERS)
γ- Coronavirus	Coronavirus de la bronchite infectieuse aviaire (IBV)
	Coronavirus du dindon (Turkey coronavirus)
Δ- coronavirus	Bulbul coronavirus HKU11, HKU12, HKU13

Cependant un tel critère de classification ne reflète que les propriétés d'un sous-ensemble limité de protéines virales surtout que depuis quelque temps des cas sont survenus où des virus clairement apparentés au groupe 1 se sont avérés ne pas induire une réaction croisée sérologique (9,15). Presque tous les virus des groupes 1 et 2 ont des hôtes mammifères, les coronavirus humains appartenant à chacun de ces groupes. Les virus du groupe 3 en revanche ont été isolés uniquement à partir d'hôtes aviaires (Tableau 2) (15).

Tableau 2 : Classification de la famille des coronavirus selon l'hôte, le virus en cause, les pathologies provoquées et le types de réponses cellulaires (9)

Group	Hôtes	Virus	Maladies	Réponses cellulaires
I.(Pathogènes Animaux)	Humain	229E et NL-63	infections Respiratoires	APN et ACE2 Humain
	porc	TGEV, PRCoV (Porcine coronavirus)	infections Respiratoires et entériques	Porcine APN
	Chat	Canine coronavirus, FeCoV (féline coronavirus), FIPV	Infections Respiratoires , entériques, neurologiques , et des hépatites	Canine et féline APN
II. (Pathogènes vétérinaires)	Humain	OC43, HKU1, et SARS-CoV	infections Respiratoires, et possibles infections entériques	contient un groupement Neu5,9Ac2-
	Souris	MHV(Murine Hepatite)	Infections entériques, neurologiques , et des hépatites	Murine CEACAM1
	Rat	Sialodacryoadenitis coronavirus	Infections neurologiques	Pas encore déterminé
	porc	Hemagglutinating encephalomyocarditis	Infections Respiratoires , entériques, neurologiques , et des hépatites	Neu5,9Ac2- contient un groupement
	bovin	BCoV	Infections entériques	Neu5,9Ac2- contient un groupement
III.(pathogènes aviaires s)	Turque	Turkey coronavirus	infections Respiratoires et entériques	Pas encore déterminé
	Poulet	IBV	infections Respiratoires et des hépatites	Pas encore déterminé

1. APN : Amino Peptidase N

2. ACE2 : enzyme de conversion 2

3. Neu5,9Ac2 : Acide N-acétylneuraminique

4. IBV : virus d la bronchite infectieuse

5. TGEV : Virus de la gastro-entérite transmissible.

6. FIPV : virus de péritonite infectieuse féline

7. CEACAM1 : « carcinoembryonic antigen cell adhesion molecule 1 »

1.3 Historique des coronavirus

En 1930 les scientifiques ont découvert le premier coronavirus qui est un virus animal responsable chez les volailles d'une maladie respiratoire aiguë très contagieuse et désigné sous le nom de virus « IBV » (virus de la bronchite infectieuse aviaire). Puis en 1946, le virus de la gastroentérite du porc « TGEV » (*Transmissible Gastroenteritis Virus*) est identifié, et en 1951 une équipe londonienne découvre le virus de l'hépatite murine « MHV » (*Murine Hepatitis Virus*) chez une souris paralysée (16).

Le premier coronavirus humain est découvert en 1965 dans les prélèvements respiratoires d'un écolier présentant un rhume banal par des chercheurs britanniques qui l'ont baptisée B814. L'année suivante en 1966 un deuxième coronavirus humain est découvert, nommé HCoV 229E, il ressemble sur le plan morphologique au coronavirus humain B814 ainsi qu'à celui de la bronchite infectieuse aviaire. Le matériel génétique du virus HCoV 229E est composé d'ARN et devient une « souche prototype » pour toute une série de travaux ultérieurs. en 1967 une autre souche nommée le HCoV OC43 est découverte aux Etats-Unis par Kenneth McIntosh (17).

Vers la fin des années 1960 les scientifiques s'aperçoivent que les virus de la bronchite infectieuse aviaire, de la gastroentérite du porc et de l'hépatite murine ont la même morphologie caractéristique « en couronne » en microscopie électronique et en novembre 1968 le terme « coronavirus » fait officiellement son apparition dans la revue *Nature* (Figure 3)(15).

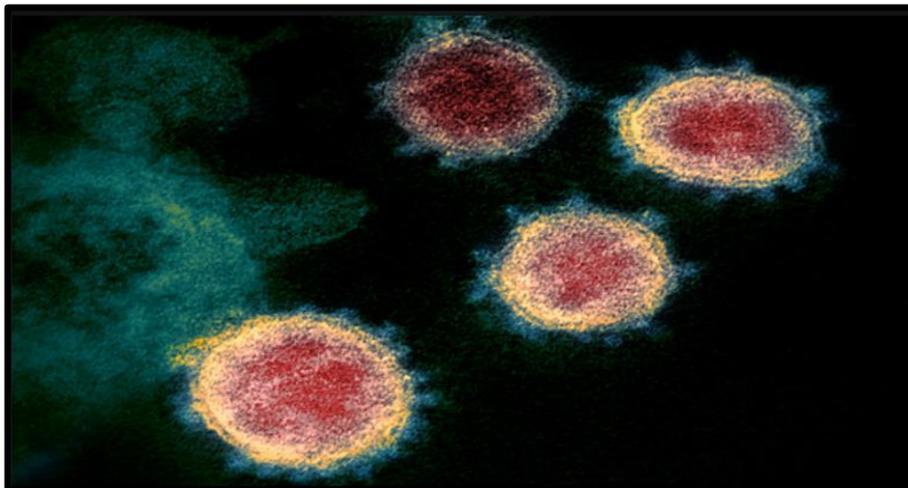


Figure 3 : Morphologie des coronavirus au microscope électronique (18).

Jusqu'au début des années 2000 les coronavirus étaient considérés comme des virus inoffensifs pour l'homme, cependant cette image qu'avaient les scientifiques sur ces virus a changé durant ces deux dernières décennies car trois nouveaux coronavirus humains ont provoqué des pandémies mondiales dangereuses pour l'humanité (15).

En novembre 2002 un nouveau coronavirus humain est découvert, il est responsable d'une maladie respiratoire très contagieuse qui aboutit en une pandémie mondiale touchant 29 pays. Les chercheurs le baptisent SARS-COV (syndrome respiratoire aigu sévère) et suggère qu'il s'est transmis de la civette palmiste masquée (*Paguna larvata*) à l'Homme. Ils ont par la suite démontré que la chauve-souris rhinolophe fer-à-cheval représente le réservoir naturel de ce virus (8).

En 2012 une deuxième épidémie mondiale se déclare, elle débute au moyen orient puis se propage et touche quelques pays. Elle est due à un nouveau coronavirus humain nommé HCOV-MERS, ce virus se serait d'abord transmis du dromadaire (réservoir naturel) vers l'homme puis se serait propagé par transmission interhumaine provoquant ainsi une épidémie dans le monde (19).

2 Généralité sur la COVID19

2.1 SARS-COV-2

2.1.1 Définition de la COVID19 « SARS-COV-2 »

La **COVID19** ; "CO" signifie corona, "VI" virus et "D" a été choisi pour "disease" (maladie de Coronavirus 2019) est une maladie causée par un virus appartenant à la famille des Coronaviridae, sous-famille Orthocoronavirinae, genre Bétacoronavirus, sous genre Sarbecovirus (acronyme pour SARS-like Bétacoronavirus), nommé **SARS-COV-2** « Syndrome Aigue Respiratoire Sévère ». c'est un virus enveloppé et sphérique dont le génome est un ARN de polarité positive directement traduisible en protéines (Figure 4)(20).

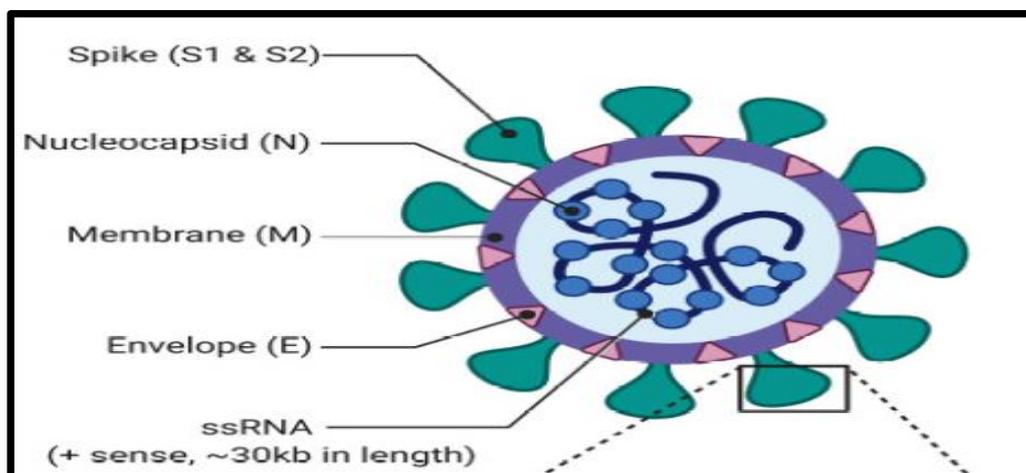


Figure 4 : Aspect structurelle des coronavirus (19)

2.1.2 Historique de la COVID19

Fin 2019 un groupe de patients atteints de pneumonie de cause non identifiée est apparu à Wuhan, dans la province de Hubei, en Chine(6). Cette nouvelle maladie est due à un nouveau coronavirus humain, initialement appelé COVID-19 il a été signalé à l’OMS (Organisation mondiale de la santé) le 31 décembre 2019. En peu de temps l’infection par ce virus s’est propagé et a été signalé dans d’autres pays et en moins de quelques mois une pandémie s’est produite et de nombreux pays étaient atteints(21).

En chine après l’augmentation des nombres de cas atteints et devant l’incapacité des médecins à diagnostiquer la maladie et à trouver un traitement les communautés chinoises de santé publique ont pris des mesures pour permettre une reconnaissance rapide de ce nouveau coronavirus humain. Plusieurs jours après, des patients similaires sont apparus dans d’autres villes de la Chine, dont la plupart avaient des antécédents d’exposition à Wuhan dans les 14 jours précédents. Le 22 janvier la province d’Anhui a confirmé un premier cas (21,22).

Le 11 février 2020, l’OMS a donné à la nouvelle infection à coronavirus son nom officiel « **COVID-19** » et le Comité international de taxonomie des virus a nommé l’agent infectieux **SARS-COV-2**. Actuellement la pandémie a touché les sept continents (Figure 5) plus particulièrement trois continents ; l’Europe, l’Amérique et l’Asie (23).

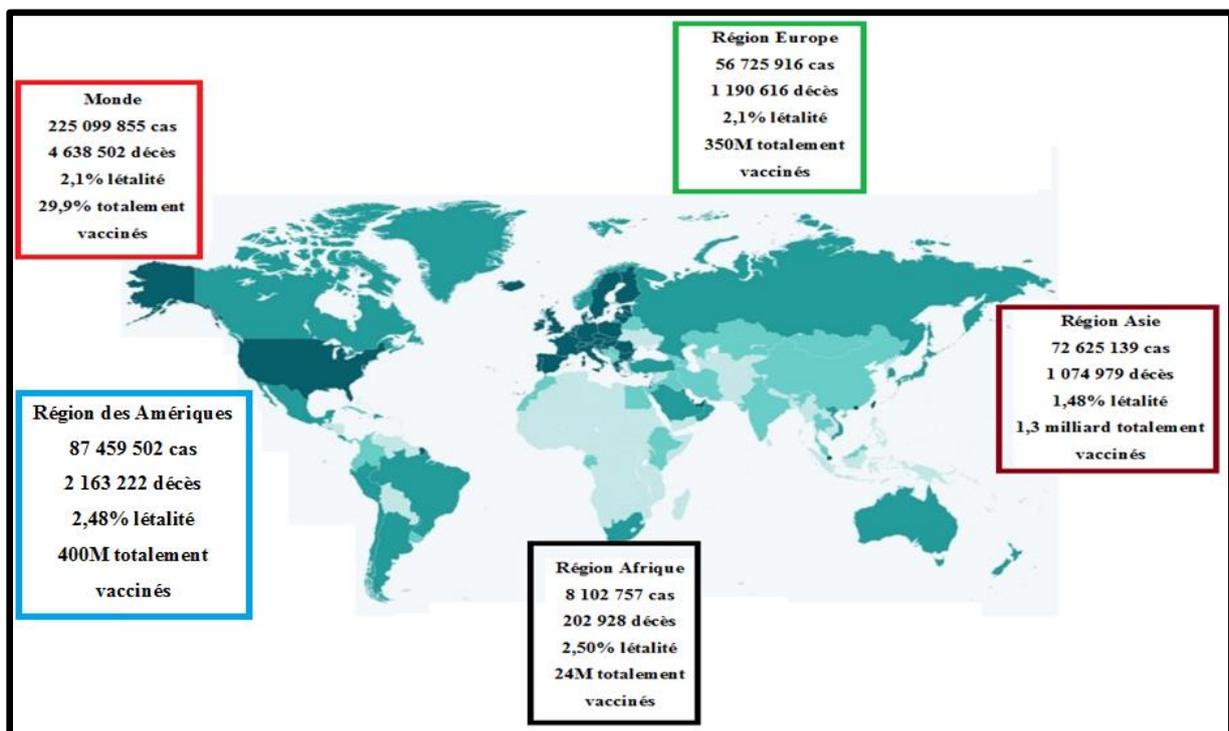


Figure 5 : distribution du SARS-COV-2 dans le monde (source WHO).

Selon le dernier rapport de l'OMS, les pays les plus touchés sont : États-Unis, Inde, Brésil, France, Russie, Espagne, Royaume-Uni, Turquie, Argentine, Iran et Colombie (17).

Le continent Africain a faussé toutes les prédictions des scientifiques, bien que tous les pays soient atteints, la pandémie n'a pas fait autant de ravage que dans les autres continents car l'Afrique a été confrontée à de nombreuses épidémies tout au long de son histoire (24).

Le premier cas confirmé du SARS-COV-2 en Afrique a été signalé en Égypte le 14 février 2020 et le deuxième un jour plus tard en Algérie. En mars, des cas de SARS-COV-2 étaient signalés dans la majeure partie du continent. Comme dans le reste du monde les premiers cas ont été importés d'autres régions suivis d'une transmission locale. Début avril presque tous les pays d'Afrique étaient atteints avec des centaines à des milliers de cas signalés dans les pays les plus durement touchés (25).

L'Algérie étant le deuxième pays touché après l'Égypte par la pandémie, elle se propage à partir du 25 février 2020 lorsqu'un ressortissant Italien est testé positif au **SARS-COV-2** puis le 1^{er} mars 2020 un foyer de contamination se forme à Blida puis se propage pour toucher toutes les wilayas algériennes, mais grâce à des mesures prises par le gouvernement ; mise en quarantaine des wilaya touchées, instauration d'un couvre-feu , fermeture des institutions scolaires, fermetures totales des frontières, campagnes instructives pour le respect des gestes barrières et fermetures des commerces non essentiels , le gouvernement algérien a réussi à garder la pandémie sous contrôle (25).

Le 11 mars 2020, l'Organisation mondiale de la santé (OMS) a déclaré l'état d'urgence sanitaire et l'apparition d'une nouvelle pandémie mondiale (22).

2.1.3 Épidémiologie du SARS-COV-2

En général l'émergence d'une maladie infectieuse dépend de trois éléments essentiels : la source de l'infection, la voie de transmission et la population sensible. En ce moment les patients infectés par le **SARS-COV-2** sont la principale source d'infection via la voie respiratoire (des gouttelettes émises par la toux , l'éternuement et la salive) (26).

2.1.3.1 Distribution géographique du SARS-COV-2

Depuis l'apparition du SARS-COV-2 en Chine, le nombre des cas infectés a très vite augmenté et après trois mois a dépassé déjà les 80000 cas (Figure 6) , mais grâce à la politique et à la réponse rapide du gouvernement chinois et aux efforts fournis par son peuples la situation est actuellement stable(23), Malgré la victoire de la Chine dans la lutte contre le SARS-COV-

2 la situation à l'extérieur été devenue plus sérieuse, plus de 160 pays de six continents avaient signalé des cas de SARS-COV-2, plus de 200000 cas signalés en Italie, l'Espagne, l'Allemagne et les États-Unis d'Amérique (avril 2020, OMS)(20).

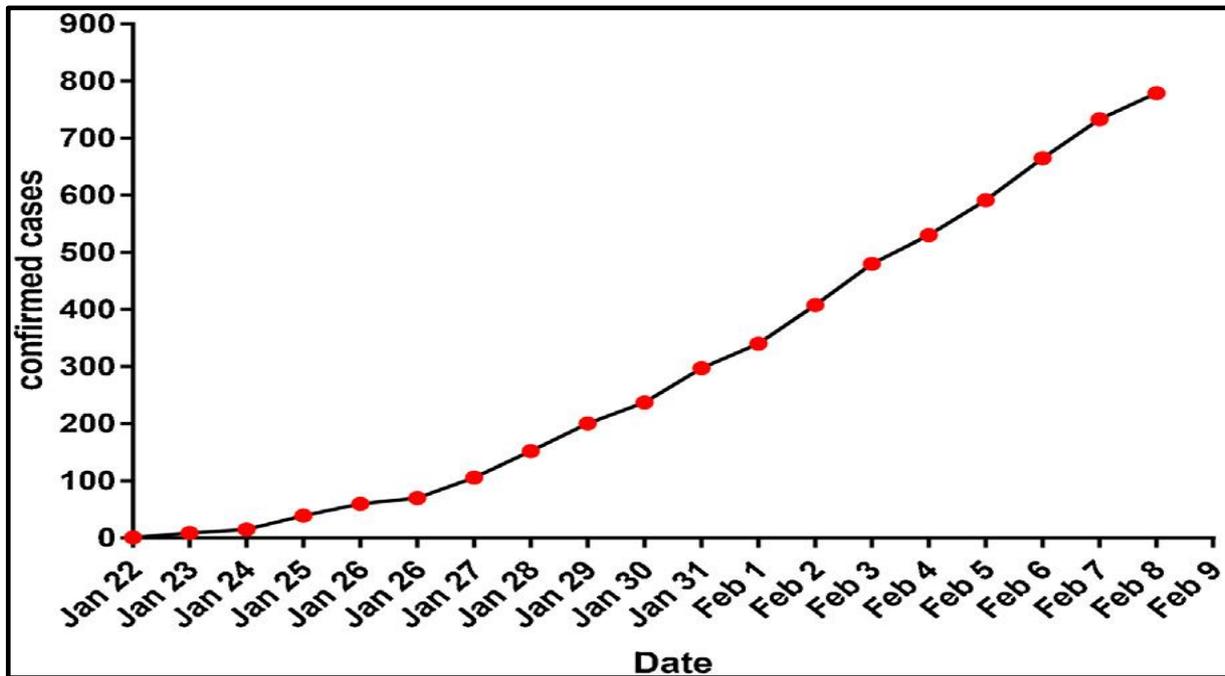


Figure 6 ; évolution de la pandémie dans la province de Anhui, Chine (23).

Les autres pays du continent asiatique n'ont pas connu de situation de crise sanitaire grave et ont bien géré la pandémie, sauf le Japon qui a atteint les 1,5 millions de cas et environ 16 400 décès et l'Inde où la situation sanitaire est grave, c'est le deuxième pays le plus touché au monde avec plus de 33 millions de cas et 441 443 décès(6,23).

Les États-Unis sont de loin le pays de plus touché au monde avec au total 41 millions de cas confirmés et 669 022 cas de décès devant l'Inde avec 33 millions de cas confirmés et 441 443 cas de décès et le Brésil 20 millions et 584 208 cas de décès (OMS)(6).

Sur le continent Africain, l'Afrique du Sud est le pays le plus atteint avec 2,8 millions de cas et 83 899 décès, le Maroc et la Tunisie complètent le podium à la deuxième et troisième place (Figure 7) (Source OMS).

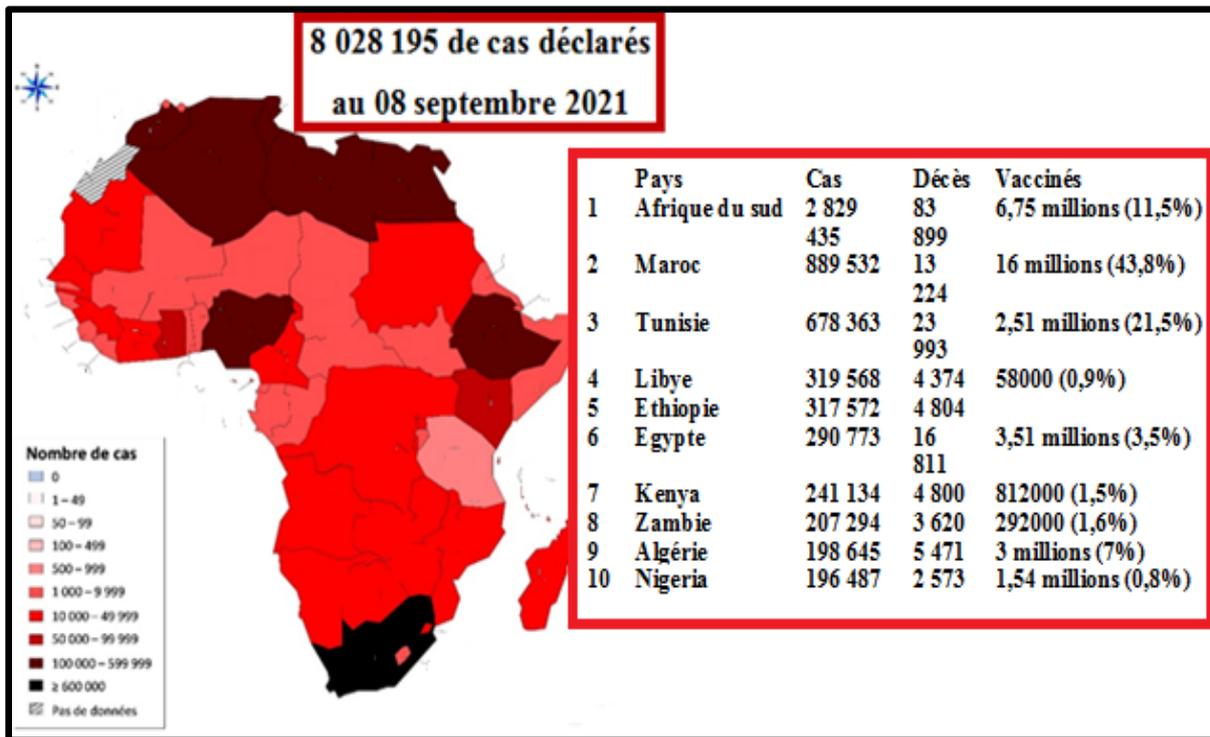


Figure 7 : Cas confirmés du SARS-COV-2 déclarés en Afrique au 08 septembre 2021 (OMS)

A l'heure actuelle l'Algérie est le neuvième pays d'Afrique le plus touché, elle compte environ 198 645 cas confirmés dont 135000 cas guéris et environ 5400 cas de décès (Figure 8), à titre de comparaison ; le Maroc compte 889 532 cas confirmé et 13 224 cas de décès, la Tunisie 678 363 cas confirmé et 23 993 cas de décès (Source OMS).

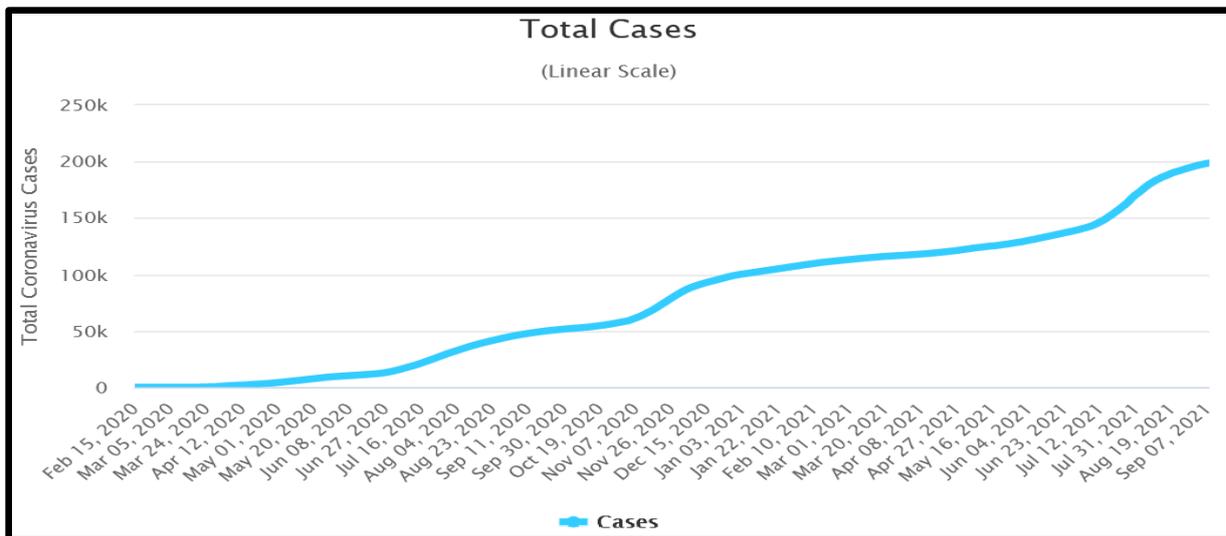


Figure 8 : Évolution quotidienne des cas de SARS-COV-2 déclarés en Algérie entre le 15 février 2020 et le 07 septembre 2021 (OMS).

A ce jour 222 893 534 cas confirmés de SARS-COV-2 ont été signalés dans le monde et ce malgré les mesures prises par les pays et les campagnes de vaccinations (Source OMS).

Tableau 3 : épidémiologie du SARS-COV-2 selon les continents et leurs pays les plus touchés (21).

Les cinq pays les plus atteints par continent			
Continents	Pays	Cas	Décès
Asie	Inde	33 millions	441 443
	Turquie	6,5 millions	58 651
	Iran	5,2 millions	112 431
	Indonésie	4,1 millions	137 782
	philippine	2,1 millions	34 672
Europe	Russie	7 millions	189 582
	Royaume-Uni	7 millions	133 483
	France	6,8 millions	115 159
	Espagne	4,9 millions	85 066
	Italie	4,5 millions	129 638
Afrique	Afrique du sud	2,8 millions	+83 899
	Maroc	889 532	13 224
	Tunisie	678 363	23 993
	Libye	319 568	4 374
	Ethiopie	317 572	4 804
Amérique	Etats-Unis	41 millions	669 022
	Bésil	20 millions	584 208
	Argentine	5 millions	112 851
	Colombie	4,9 millions	125 378
	Mexique	3.4 millions	264 541

2.1.4 Origine du SARS-COV-2

Les animaux sauvages sont considérés comme des hôtes réservoirs naturels et sont responsables de la transmission de divers virus, dont l’Ebola, le Nipah, et les Coronavirus (27).

Comme pour le SARS-COV-1 et le MERS, le SARS-CoV-2 est probablement originaire des chauves-souris (réservoir naturel) (Tableau 4) (27).

Tableau 4 : Hôtes naturels et intermédiaires des coronavirus infectant l’homme (29).

Coronavirus	Hôte naturel	Hôte intermédiaire
HCoV-NL63	Chauves-souris	
HCoV-229E	Chauves-souris	Dromadaire
HCoV-OC43	Rongeurs	Bovins
HCoV-HKU1	Rongeurs	
SARS-COV-1	Chauves-souris	Civette palmiste masquée
SARS-COV-2	Chauves-souris	Pangolin
MERS	Chauves-souris	Dromadaire , Lama

Le SARS-COV-2 est très étroitement lié à une souche de coronavirus identifiée chez la chauves-souris le **RATG13** « *Bat Rhinolophus affinis, Tongguan, 2013* ».qui sont identiques à 97,41 % au niveau du génome (28), cependant il n’est pas encore certain que le virus ait été transmis directement par les chauves-souris et c’est dans cette perspective que l’idée de l’existence d’un hôte intermédiaire a surgi, actuellement le pangolin est considéré comme étant l’hôte intermédiaire le plus plausible, les souches isolées chez ce derniers indique une identité égale à 92,6% cependant montre une similarité plus grande au niveau du domaine **RBD** de la protéine S égale à 97,4% alors qu’elle est seulement égale à 89,2% pour le RATG13 (Figure 9) (26,29).

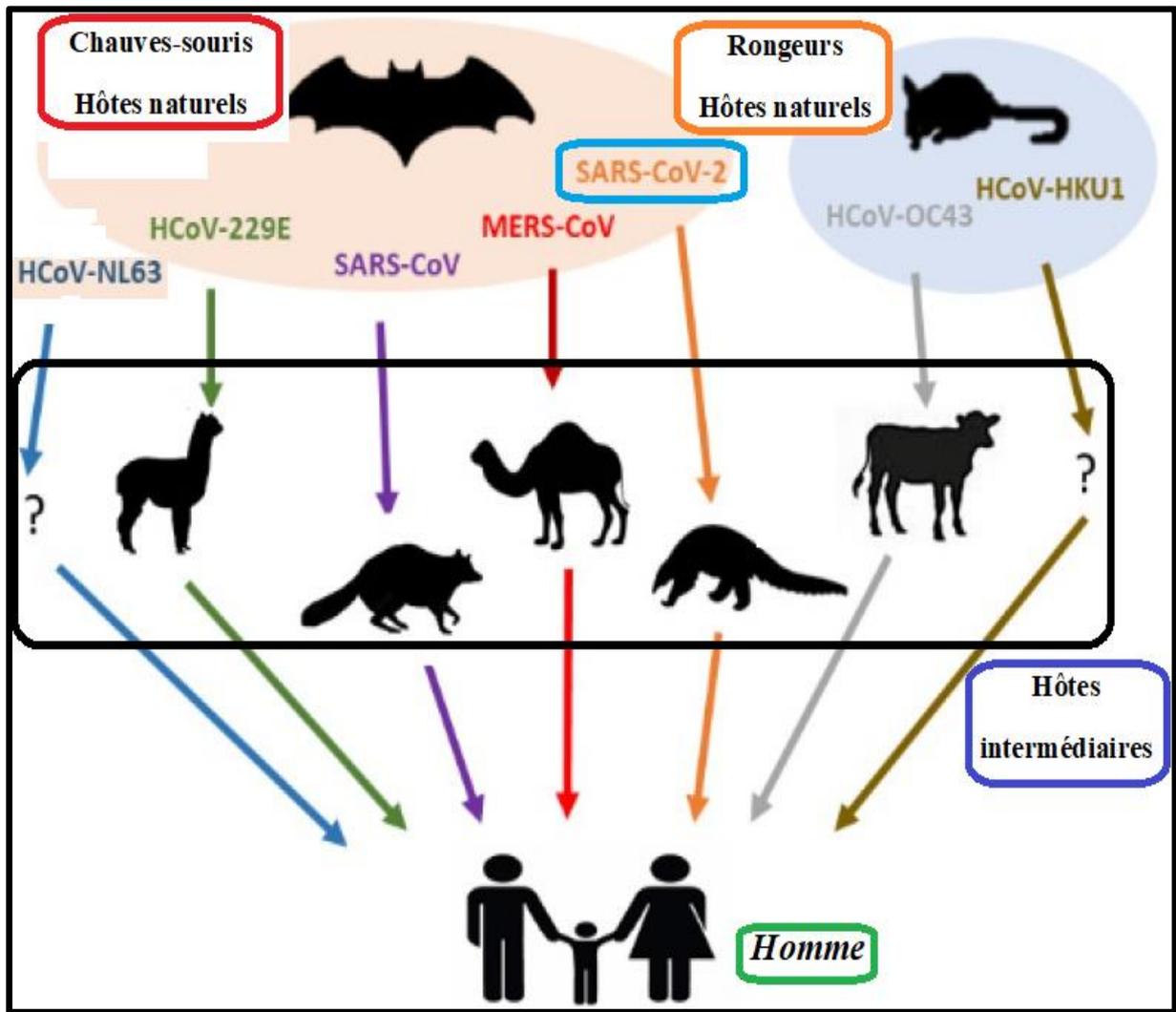


Figure 9 : les hôtes naturels et intermédiaires des Coronavirus (28).

2.2 Les variants du SARS-COV-2

2.2.1 Définitions des variants du SARS-COV-2

Les virus évoluent dans le temps par des mutations et tout particulièrement lorsqu'il y a beaucoup d'infections dans une population (6).

Les mécanismes d'évolution des virus impliquent des mutations (changements du code génétique) et des recombinaisons. Ces mutations arrivent au hasard et se produisent pendant la réplication virale à l'intérieur des cellules de l'hôte. Si le virus est porteur d'une ou plusieurs mutations par rapport au virus initial alors il est désigné sous le nom de « variants ». Cependant si le comportement de ce variant est différent (plus transmissible, plus pathogène, infectant d'autres types cellulaires) on parlera alors de « souche » virale (30).

Les mutations les plus fréquentes sont :

- Des substitutions : une base nucléique est remplacée par une autre ce qui peut entraîner une modification de l'acide aminé correspondant.
- Des délétions : une ou plusieurs bases nucléiques sont éliminées et par conséquent la perte de l'acide aminé correspondant également
- Des insertions : des triplets de bases sont introduits et donc de nouveaux acides aminés (31).

2.2.2 Nomenclature des variants du SARS-COV-2

Actuellement trois systèmes de nomenclature sont utilisés pour nommer et suivre les variants du SARS-CoV-2 : **GISAID** (*Global Initiative on Sharing Avian Influenza Data*), **Nextstrain** et **Pango**. Chaque système a sa propre approche scientifique pour classer et nommer les lignées. L'existence de différents systèmes de nomenclature peut signifier que le même variant porte plusieurs noms (Tableau 5) (32).

Tableau 5 : Nomenclature scientifique des variants du SARS-COV-2 (32).

Nomenclature	Pango	GISAID	Nextstrain	Pays
Variants	B.1.1.7	GR501Y.V1	20I/S : 501Y.V1	Royaume-Uni
	B.1.351	GH501Y.V2	20H/S : 501Y.V2	Afrique du sud
	P1	GR501Y.V3	20J/S : 501Y.V3	Brésil
	B.1.617.2	G/452R.V3	21A/S : 478K	Inde

Cependant le manque de clarté dans la dénomination des variants fait qu'il est difficile pour les non experts dans le domaine de relier ces variants aux publications scientifiques. De plus, l'utilisation de systèmes de nomenclature multiples est source de confusion pour les responsables de la santé, les médias et le public.

C'est pour ces raisons que l'organisation mondiale de la santé (OMS) a réunis des experts en nomenclature virologique et microbienne afin de trouver une nomenclature moins scientifique et plus facile à assimiler pour la population. Ainsi la nouvelle nomenclature consiste à désigner les variants par des lettres grecques (Alpha, Béta, Gamma) (Tableau 6) (3).

Tableau 6 : Nomenclature des variants du SARS-COV-2 selon l'OMS (32).

Nomenclature	OMS	Pango	GISAID	Nextstrain	Pays
Variants	Alpha	B.1.1.7	GR501Y.V1	20I/S : 501Y.V1	Royaume-Uni
	Béta	B.1.351	GH501Y.V2	20H/S : 501Y.V2	Afrique du sud
	Gamma	P1	GR501Y.V3	20J/S : 501Y.V3	Brésil
	Delta	B.1.617.2	G/452R.V3	21A/S : 478K	Inde

2.2.3 Historique des variants du SARS-COV-2

Le premier variant du SARS-COV-2 est apparu vers fin janvier ou début février 2020 en Chine et en quelques mois il a remplacé la souche initiale du SARS-CoV-2 devenant ainsi en juin 2020 le virus dominant en circulation dans le monde. Ce premier variant est désigné sous le nom de G614 en référence à la mutation qu'il porte (33).

En mai 2020 un autre variant du SARS-COV-2 est découvert en Afrique du sud, ce nouveau variant est nommé B.1.351 (PANGO) ou Béta (OMS) ou 501Y.V2 en raison de la mutation N501Y qu'il porte. Il se propage rapidement en Afrique du sud et remplace les autres souches du SARS-CoV-2 circulant dans le pays devenant ainsi le 18 décembre 2020 un variant préoccupant et est baptisé VOC-20DEC-02 (second Variant Of Concern, december 2020) (34).

En juin 2020 un variant du SARS-CoV-2 transmis par des visons d'élevage est découvert au Danemark. Ce variant appelé « Cluster 5 » par les autorités danoises présente une combinaison de mutations jamais observée auparavant (35).

Fin septembre 2020 un nouveau variant du SARS-COV-2 est apparu aux Royaume-Uni, désigné sous le nom de B.1.1.7 (PANGO) ou 501Y.V1 ou variant Alpha (OMS), il est apparu initialement dans le sud-est de l'Angleterre puis en quelques semaines a progressivement remplacé les autres souches virales existantes dans le pays et s'est propagé dans 62 pays d'Europe, d'Asie, des États-Unis et d'ailleurs (36).

En Avril 2020 le variant brésilien émerge, il est initialement nommé 501Y.V3, en décembre 2020 l'OMS le classe dans la liste des variants préoccupant et le baptise variant « P1 » (19).

En décembre 2020 le premier variant Indien est découvert et est baptisé B.1.617 (PANGO). ce variant comporte trois sous souches dont deux nommées respectivement B.1.617.1 et B.1.617.2 sont apparues en décembre 2020 et le troisième B.1.617.3 en mai 2021, actuellement la sous souche B.1.617.2 ou variant Delta est la souche dominante dans le monde (34).

2.2.4 Epidémiologie des variants du SARS-COV-2

Les variants du SARS-COV-2 sont probablement sous-déclarés par la plupart des pays, car les tests les plus couramment utilisés (RT-PCR) ne font pas la distinction entre les variants à l'exception des variants contenant la délétion H69/V70 , ainsi le séquençage de l'ARN est nécessaire pour la détection des variants (30).

2.2.4.1 Dans le monde

2.2.4.1.1 Le variant Anglais « Alpha »

Le 08 septembre 2021, le variant Alpha est signalé dans plus de 200 pays et dénombre environ 1,2 millions de cas dont 14 l'ont signalé étant l'un des variants dominants : Royaume-Uni compte 270 000 cas, Etats-Unis (218 325 cas), Allemagne (102 644 cas), Suède (63 038 cas), Danemark (62 854 cas), Japon (42 225 cas), Canada (36 372 cas), France (33 916 cas), Pays-Bas (29 841 cas), Italie (25 399 cas), Espagne (23 213 cas), Suisse (21 811 cas), Belgique (20 781 cas), Irlande (15 993 cas) (Figure 10) (Source OMS).

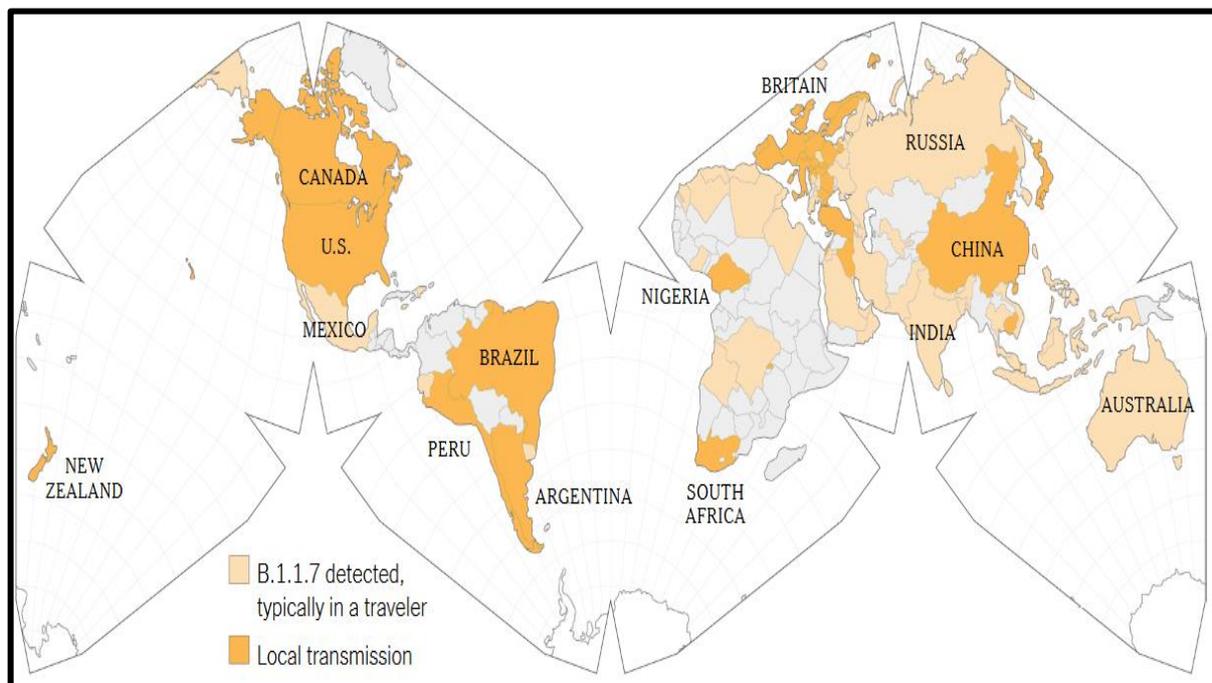


Figure 10 : Distribution géographique du variant Anglais (OMS).

2.2.4.1.2 Le variant Sud-Africain « bêta »

Le variant Bêta a été détecté pour la première fois en Mai 2020 dans la région métropolitaine de Nelson Mandela Bay dans la province du Cap-Oriental, en Afrique du Sud.

Actuellement il est détecté dans presque tous les pays du monde avec au total 28 380 cas confirmés, il reste le variant dominant en Afrique du sud avec 6154 cas (Figure 11) (OMS).

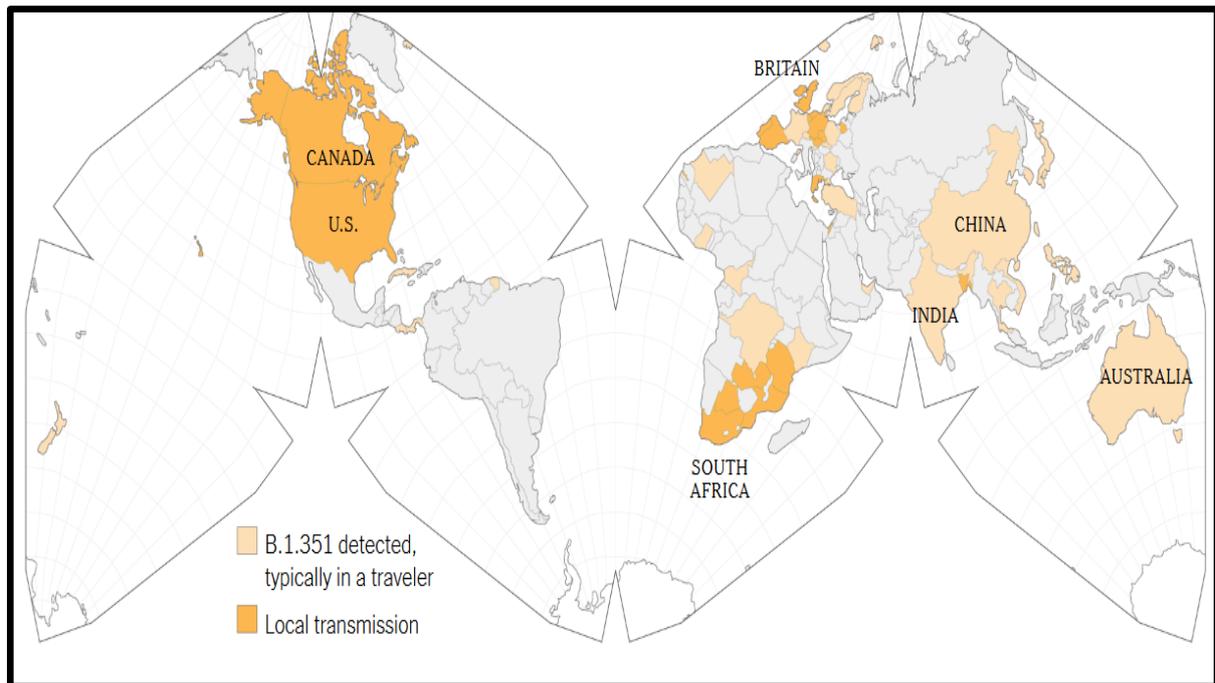


Figure 11 : Distribution géographique du variant Sud-Africain (OMS).

2.2.4.1.3 Le variant Brésilien « Gamma »

Le variant Gamma a été détectée pour la première fois dès avril 2020 dans l'amazone en Brésil. Le 6 Janvier 2021 il est détecté au japon chez quatre personnes qui étaient arrivées à Tokyo après avoir visité les amazones, au Brésil, quatre jours auparavant.

Au 08 septembre 2021, il est signalé dans 71 pays avec un total de 63 769 cas dont 47500 détecté dans 03 pays ; Etat-Uni 23373 cas, Brésil 16200 cas, Canada 8070 cas (Figure 12) (OMS).

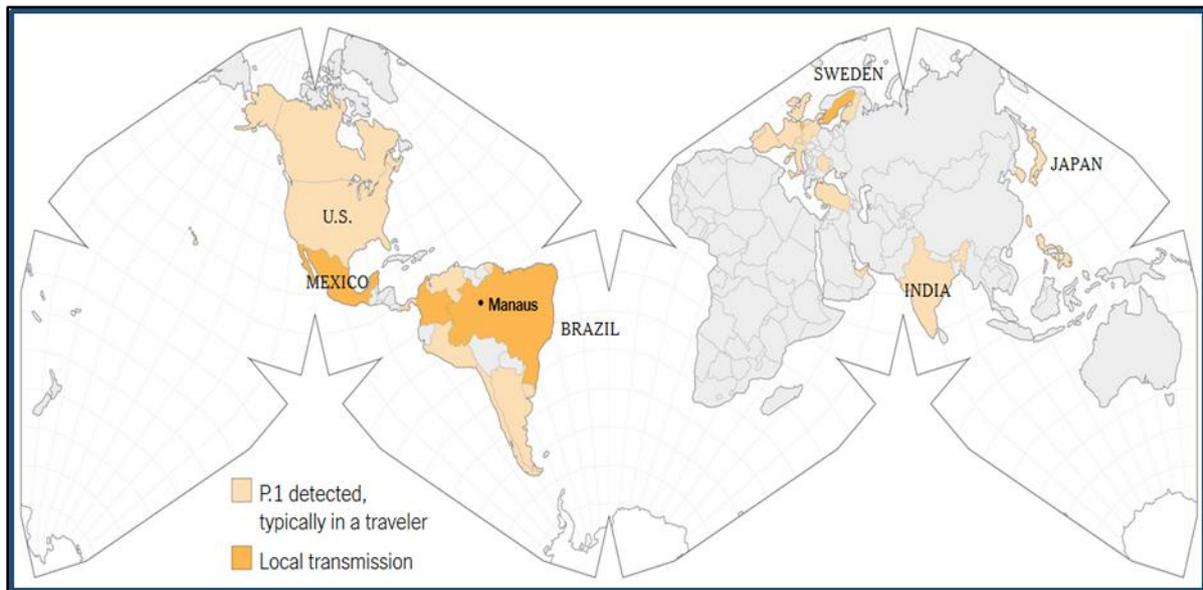


Figure 12 : Distribution géographique du variant Brésilien (OMS).

2.2.4.1.4 Le variant Indien « Delta »

le 5 octobre 2020 il est identifié pour la première fois dans le Maharashtra, en Inde. En avril 2021, il s'est propagé à au moins huit pays dont ; Angleterre, Allemagne, Canada et Singapour.

En mai 2021, les scientifiques suggèrent que le variant indien est plus contagieux que le variant britannique, ainsi au Royaume-Uni, le variant britannique n'est plus le variant dominant.

Le 08 septembre 2021, l'OMS signale le variant dans 132 pays et dénombre environ 654 296 cas dans le monde (Figure 13) (OMS).

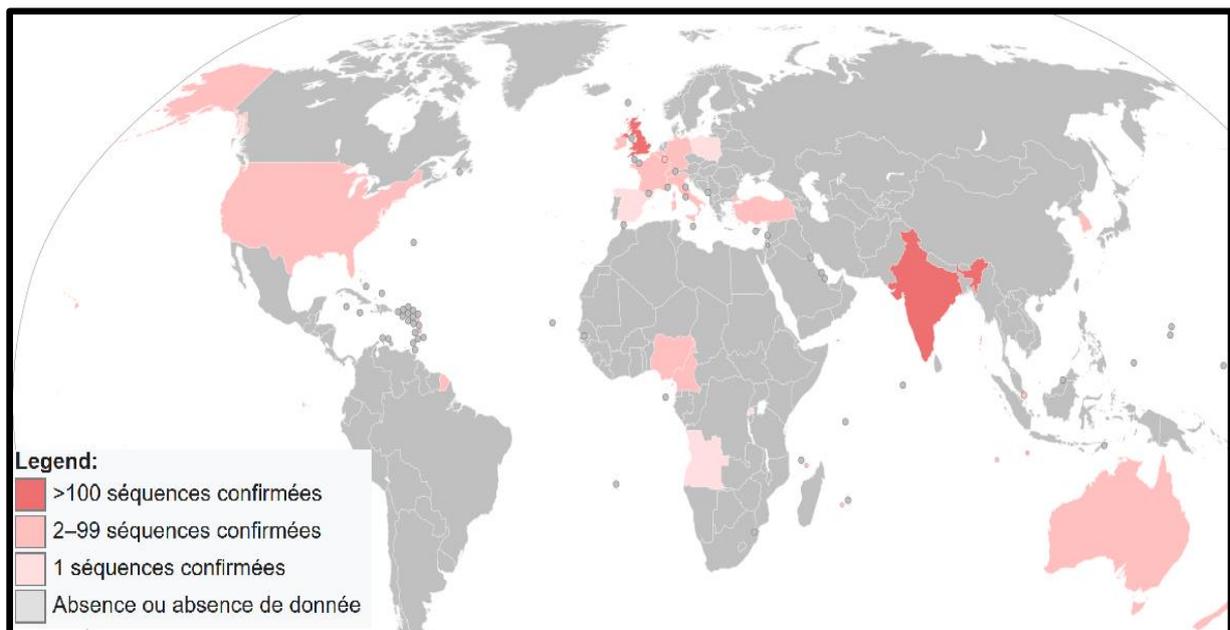


Figure 13 : Distribution géographique du variant Indien (OMS).

2.2.4.1.5 Le variant Nigérien « Eta »

Le variant Eta a été détecté en décembre 2020 au Nigeria et Royaume-Uni. Au début de février 2021 au Nigeria, le variant Eta a été identifié dans 45 % de tous les échantillons séquencés. Le 05 aout 2021, le variant Eta est signalé dans 71 pays d'Europe, d'Amérique du Nord, d'Afrique, d'Océanie et d'Asie. Les pays les plus touchés sont ; canada 1403 cas, les États-Unis d'Amérique 1184 cas, l'Allemagne 738 cas, le Royaume-Uni 517 cas, et le Nigeria 255 cas (OMS).

2.2.4.2 En Algérie

Le 12 mars 2021, 7 cas du variant Anglais sont détectés dans les wilayas d'Alger et de Blida et 13 cas du variant Nigérien dans les wilayas d'Alger, de Tébessa, de Souk Ahras, d'In Salah et de Ouargla (Institut Pasteur).

Le 31 mars, 37 cas du variant Anglais et Nigérien sont confirmés. Le 7 avril, 70 cas des variants Anglais et Nigérien sont signalés dans plusieurs wilayas. Le 14 avril, 51 nouveaux cas des variants Anglais et Nigérien dans les wilaya d'Alger, de Blida, de Djelfa, de Relizane, de Médéa, d'Oran, de Laghouat et de Touggourt selon l'Institut Pasteur d'Algérie. Le 25 avril, 143 cas dû au variant Anglais et de 230 cas dû au variant nigérien sont recensés (Institut Pasteur)

Le 3 mai, 6 cas dû au nouveau variant Indien sont détectés à Tipaza , il représentait environ 6,5% des prélèvement reçus pour séquençage contre 65% pour le variant Anglais. Cependant à partir du mois de mai le variant Indien est devenu le variant dominant et représente au mois d'aout 91% des prélèvements (données institut pasteur d'Algérie) (Figure 14) (Institut Pasteur).

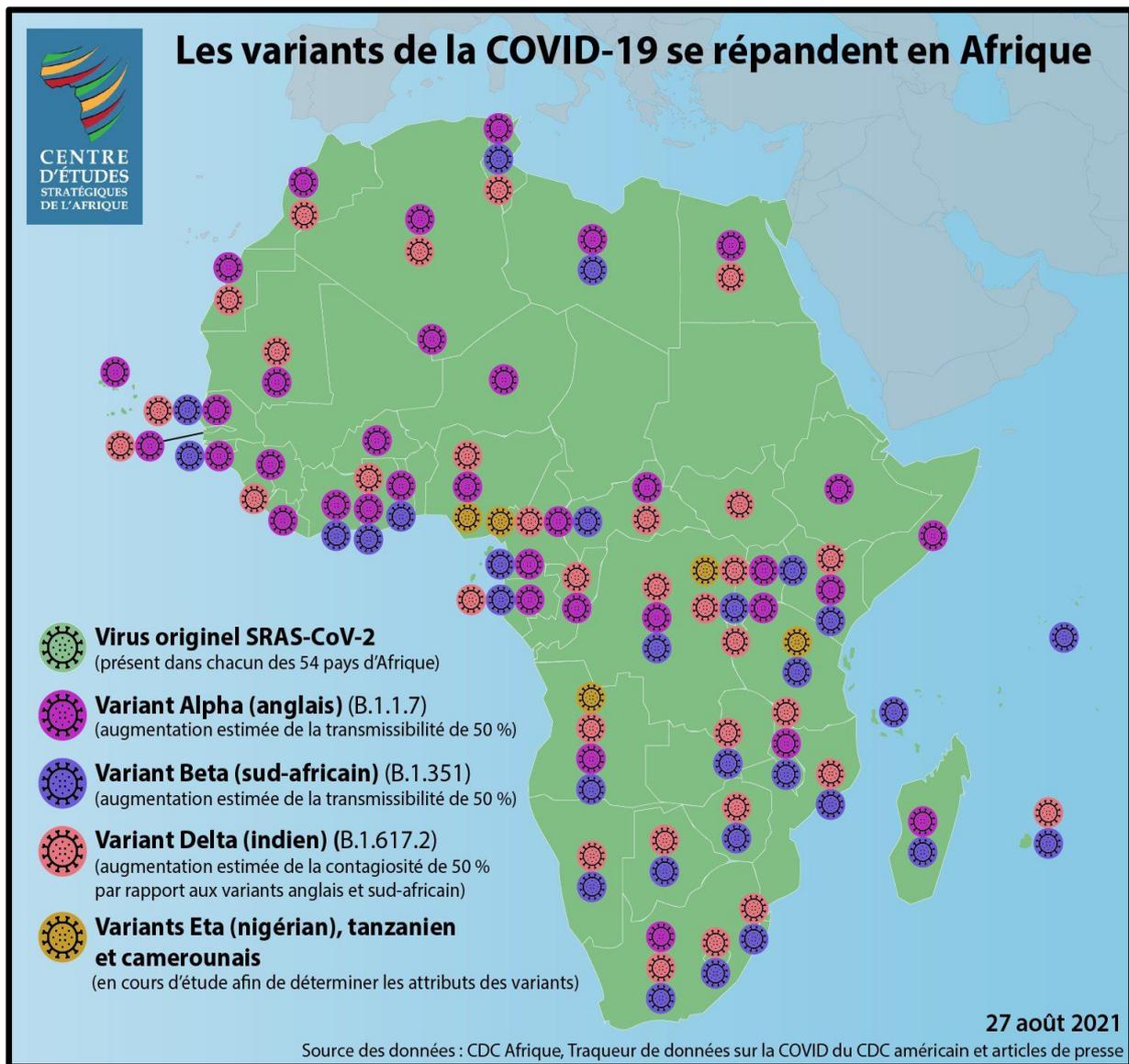


Figure 14 : Les variants détectés en Afrique et en Algérie au 27 août 2021 (CDC, Afrique).

2.3 Contagiosité du SARS-COV-2 et des variants

Le SARS-COV-2 se transmet principalement par voie respiratoire avec un taux de reproduction « R_0 » initialement estimé entre 2,5 et 3 par jour ce qui signifie que théoriquement un seul individu infecté peut contaminer en moyenne 2 à 3 personnes (Figure 15) (28) mais actuellement a été réévalué à la hausse entre 3,9 à 8 qui est vraisemblablement dû à l'apparition des variants, par conséquent il est le coronavirus le plus contagieux à ce jour en comparaison avec le SARS-COV-1 et MERS dont le R_0 est compris respectivement entre « 2,1 – 2,9 » et « 0,69 – 1,3 » (Tableau 7) (26).

Quelques cas de transmission « homme-animal » ont été observés, à titre d'exemple les animaux de compagnies : chien, chat et les animaux de zoo ; tigre, lion (37).

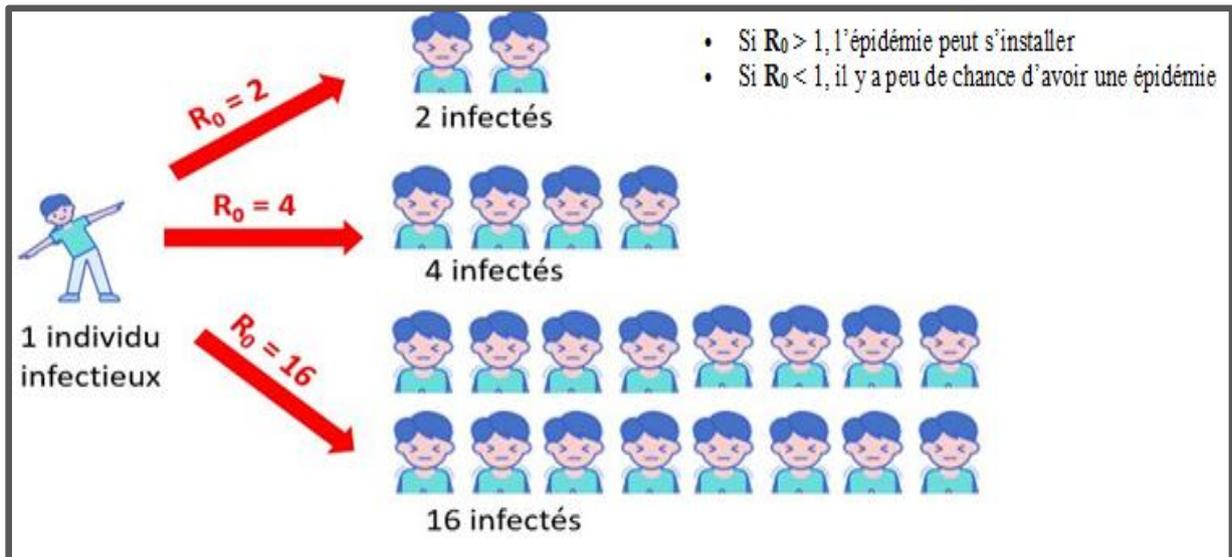


Figure 15 : Schéma expliquant le taux de reproduction de base R_0 (38).

Actuellement les variants dominants à travers le monde semble posséder un taux de reproduction R_0 plus élevé que celui du SARS-COV-2 ce qui peut suggérer une plus haute contagiosité des variants (Figure 16)(3).

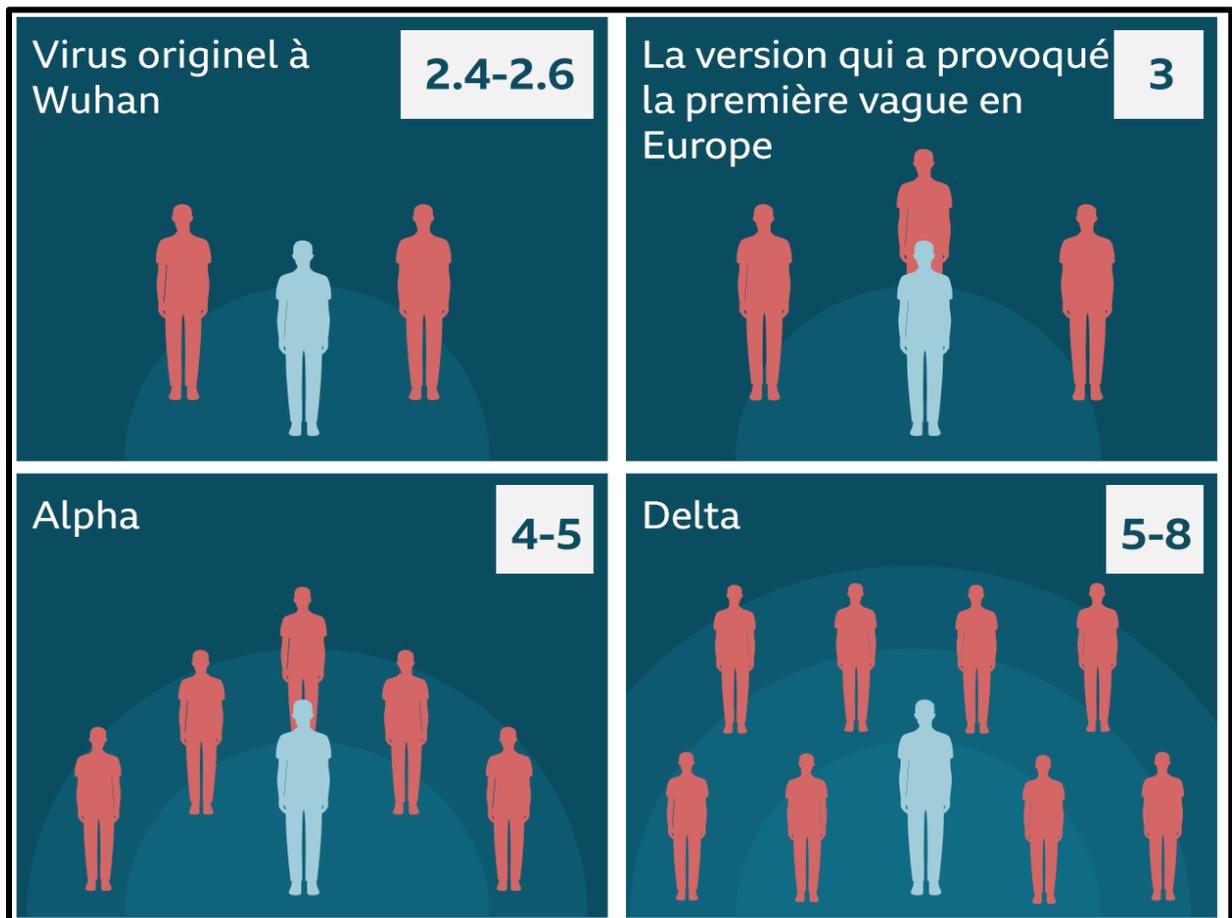


Figure 16 : Comparaison du R_0 des variants avec la souche originelle du SARS-COV-2 (39).

2.4 Taux de létalité du SARS-COV-2 et des variants

C'est le deuxième facteur épidémiologique clé, il représente la probabilité de mourir pour une personne atteinte, actuellement le taux de létalité est estimé à 2% à l'échelle mondiale (Tableau 7) mais varie en fonction de plusieurs facteurs (40,41) (Figure 17) ;

- Le virus lui-même (virulence, mutation, dose infectante)
- La population (statut immunitaire, âge, comorbidité,).
- Le système de santé (équipement, capacité d'accueil, efficacité des traitements)(24).



Figure 17 : Taux de létalité selon l'âge, le sexe et les conditions de santé (Source BBC).

2.4.1 Virulence et variants du SARS-COV-2

Depuis le début de la pandémie SARS-COV-2 en 2019 plusieurs variants du SARS-COV-2 ayant en commun les mêmes mutations sont apparus dans plusieurs pays tels que ; le variant Anglais (Alpha), le variant sud-africain (Béta), le variant Brésilien (Gamma) et le variant indien (Delta). Des études ont montré que les variants tendent à se propager plus vite car plus transmissibles ou plus infectieux, mais ne semblent pas entraîner une maladie plus grave ou une mortalité plus élevée ou des symptômes différents de ceux observés depuis le début de la pandémie (19).

2.4.2 Sexe

Les hommes présentent à la fois une gravité et un taux de mortalité plus élevés que les femmes lors d'infection par la SARS-COV-2, cette différence est principalement due au fait que les hommes ont une expression de récepteurs ACE2 dans les poumons, essentiels à l'entrée et la multiplication du SARS-COV-2 dans l'organisme, plus élevée que chez les femmes (42).

En outre cette disparité dans les statistiques peut également être attribuée à des différences de réponses immunitaires selon le sexe. La première différence réside dans le fait que les femmes possèdent deux chromosomes X tandis que les hommes ne possèdent qu'un seul chromosome X et un chromosome Y qui est plus petit et plus pauvre en gènes. Ainsi pour compenser ce déséquilibre dans la quantité de gènes l'un des chromosomes X est inactivé chez la femme pendant l'embryogénèse. Cependant le gène ACE2 qui est situé sur le chromosome Xp22.2 évite l'inactivation qui résulte en un polymorphisme plus accrue des récepteurs ACE2 chez la femme, paradoxalement l'échappement de l'ACE2 à l'inactivation entraîne une faible expression des récepteurs ACE2 au niveau des poumons. Parallèlement le chromosome X code certains gènes liés aux réponses immunitaires ce qui explique la raison pour laquelle les femmes ont une charge virale plus faible, une inflammation moins importante que les hommes et une activité plus élevée des cellules immunitaires (43).

La deuxième différence est que les hormones sexuelles féminines, les œstrogènes et la progestérone, jouent un rôle anti-inflammatoire, influencent les cellules immunitaires, stimulent la production d'anticorps, favorisent la réparation de certaines cellules respiratoires et inhibent le récepteur ACE2 et par conséquent ces hormones protègent les femmes contre certains symptômes(42,43)(Figure 18).

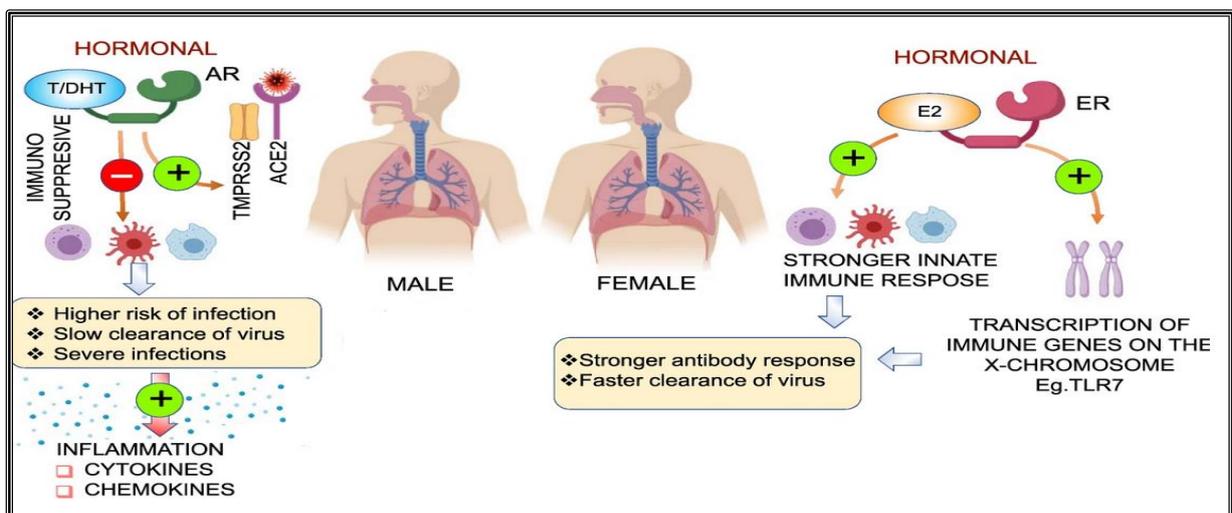


Figure 18 : variation de la réponse immunitaire selon le sexe (44).

2.4.3 Age

La gravité et le pronostic du SARS-COV-2 dépendent largement de l'âge du patient. Les personnes âgées de plus de 65 ans représentent 80 % des hospitalisations et ont un risque de décès 23 fois plus élevé que les moins de 65 ans alors qu'il avoisine 0 % chez les enfants (45).

Cet écart statistique entre ces deux tranches d'âge est d'une part dû au vieillissement du système immunitaire qui entraîne deux changements majeurs. Le premier changement est un déclin progressif de la fonction immunitaire appelé immun-sénescence qui entrave la reconnaissance, la signalisation et l'élimination des agents pathogènes. Le deuxième changement est l'augmentation chronique de l'inflammation systémique qui est due à une hyperactivité du système immunitaire mais inefficace.(46)

D'autre part la gravité de la SARS-COV-2 est fortement associée à des comorbidités telles que l'hypertension, le diabète, l'obésité, les maladies cardiovasculaires et les maladies du système respiratoire (47).

Tableau 7 : comparaison du taux de reproduction et létalité du SARS-COV-2 avec les virus des précédentes pandémies (Source OMS).

Virus	Taux de reproduction R0	Taux de létalité
Rougeole	12 – 18	0,1% - 0,7%
Variants SARS-COV-2	4 – 8	2,2 – 4,8%
Variolo	5 – 7	1% - 30%
Polio	2 – 8	5% - 15%
SARS-COV-2	2,5 – 5,5	2% - 4%
Grippe	2 – 3	0,1 %– 2,5%
SARS-COV-1	1 – 3	9,6%
Ebola	1 – 4	50%
H1N1	1 – 2	0,7%
MERS	0,6 – 1	34%

3 Biologie du SARS-COV-2

3.1 Structure et organisation génomique des coronavirus

Les coronavirus ont l'un des plus grands génomes connus parmi les virus à ARN, allant de 27 à 32 kb de longueur(48), ils sont arrondis, enveloppés, brin positif et non segmentés avec une taille d'environ 80 à 125 nm et sont entourés d'une enveloppe externe désignée sous le nom d'enveloppe virale qui est dérivée de la cellule hôte et contenant des pics de glycoprotéine appelés **Spike** ou protéine S. Le génome viral est protégé par la nucléocapside qui est de forme hélicoïdale lorsqu'elle est détendue et sphérique à l'intérieur du virus(49). Les deux tiers de la région 5' du génome codent pour seize PNS « protéines non structurales » tandis que le tiers restant code pour des protéines structurales Spike (S), Enveloppe (E), Membrane (M), Nucléocapside (N).et Hémagglutinine d'estérase HE (50). (Figure 19).

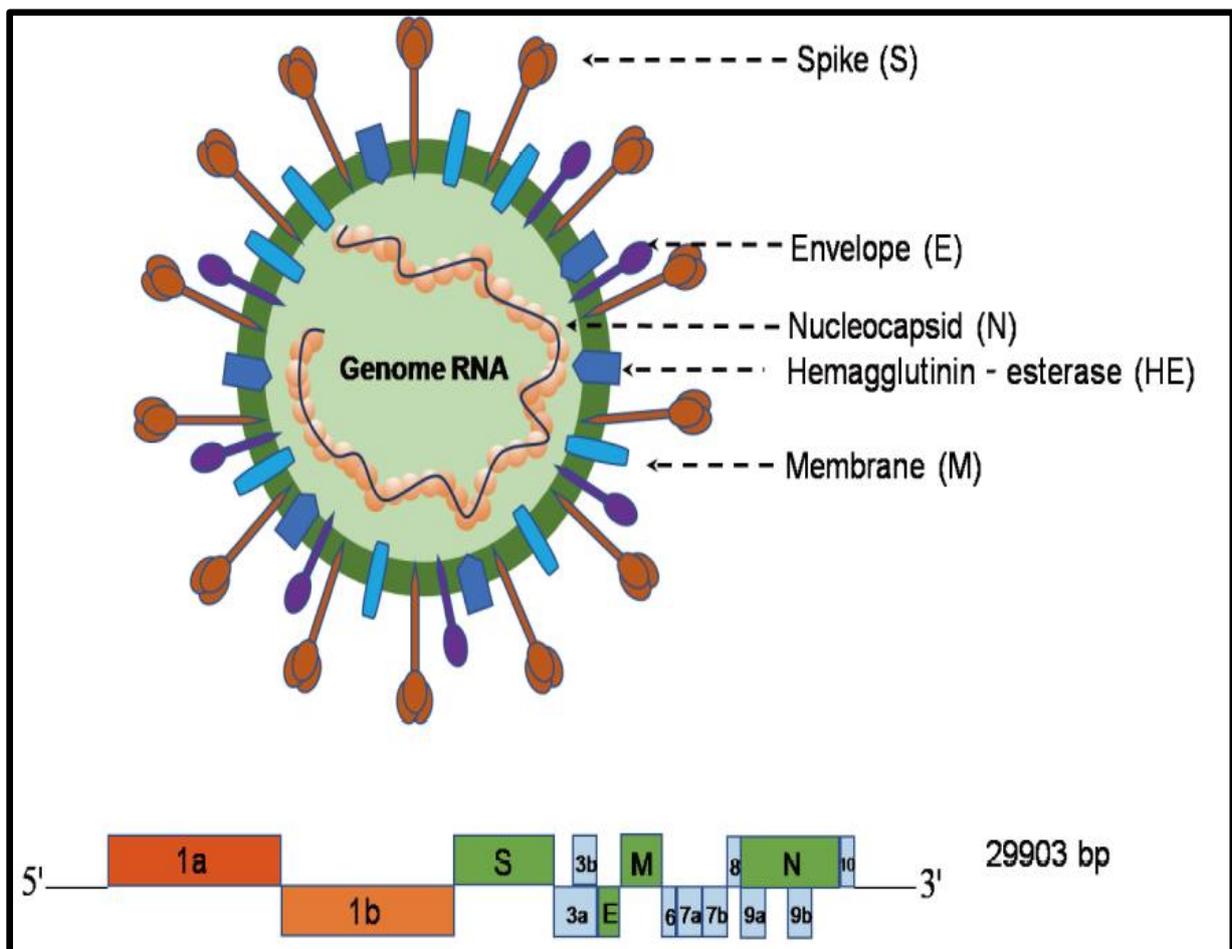


Figure 19 ; Structure et organisation du génome des coronavirus(6).

Le **SARS-CoV-2**, **SARS COV**, **HCoV-229E** et **HCoV-NL63** ont quatre gènes qui expriment les protéines de structure **S**, **M**, **N** et **E**. Les **HCoV-OC43** et **HCoV-HKU1** coronavirus ont un gène supplémentaire qui exprime la Protéine **HE** (Figure 20)(48,49).

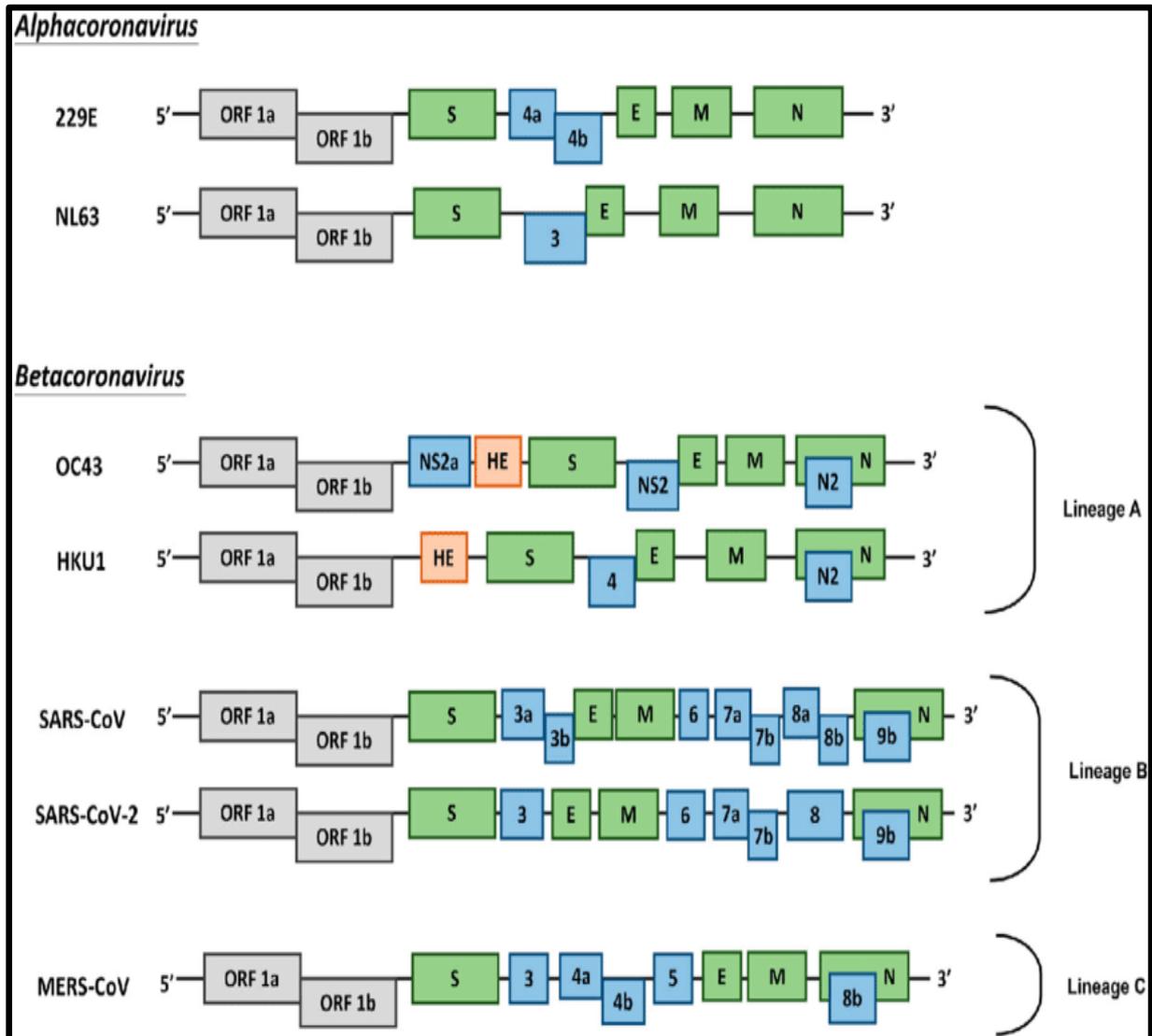


Figure 20 : comparaison de l'organisation génomiques des coronavirus (51).

3.2 Organisation génomique du SARS-COV-2

L'analyse phylogénétique a révélé que le SARS-CoV-2 est étroitement lié aux bêta-coronavirus. Comme pour les autres coronavirus, le génome du SARS-CoV-2 est un ARN simple brin positif avec une coiffe 5' -UTR, et une queue poly(A) 3'-UTR. La longueur du génome du SARS-CoV-2 est inférieure à 30 kb, il comprend quatorze ORF (Open Read Frame « cadres de lecture ouverts ») codant pour des protéines non structurales qui interviennent dans les processus de répllication et d'assemblage du virus (Figure 21) (52).

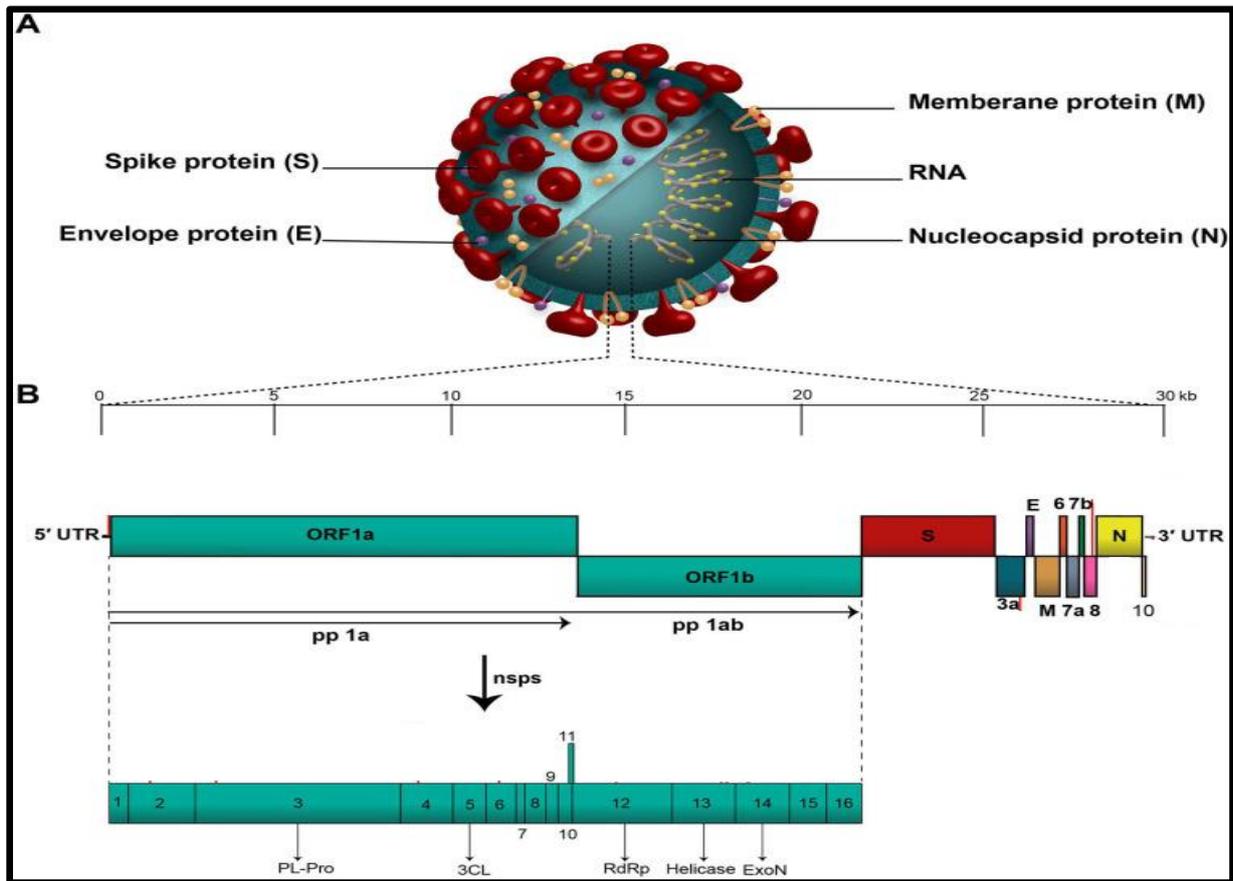


Figure 21 : Organisation du génome du SARS-CoV2 (53).

Le séquençage du SARS-COV-2 a révélé une similitude génétique avec les coronavirus précédemment connus tels que le **RATG13** chez la chauve-souris à 96,75%, le **SARS-COV-1** à 86% et le **MERS** à 81% (Figure 22) (54).

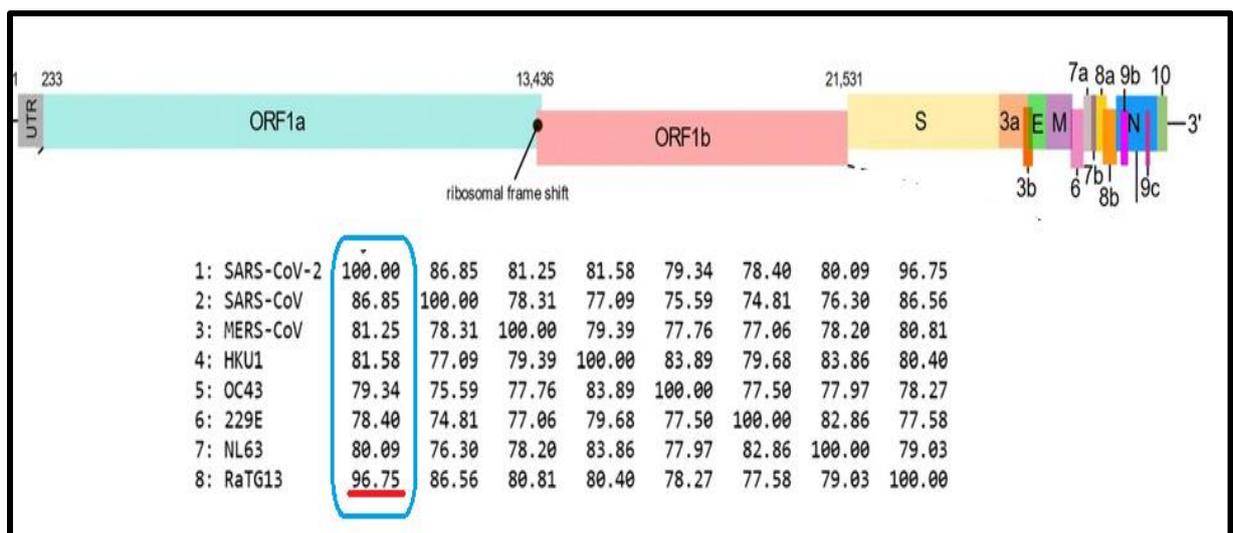


Figure 22 : pourcentage de similarité du génome du SARS-COV-2 avec les autres coronavirus (55).

3.3 Structure des protéines du SARS-COV-2

3.3.1 Protéine Spike « S »

La protéine S est une protéine transmembranaire qui facilite la liaison de l'enveloppe virale aux récepteurs de l'enzyme de conversion de l'angiotensine 2 « ACE2 » situés à la surface des cellules hôtes (52) (Figure 23).

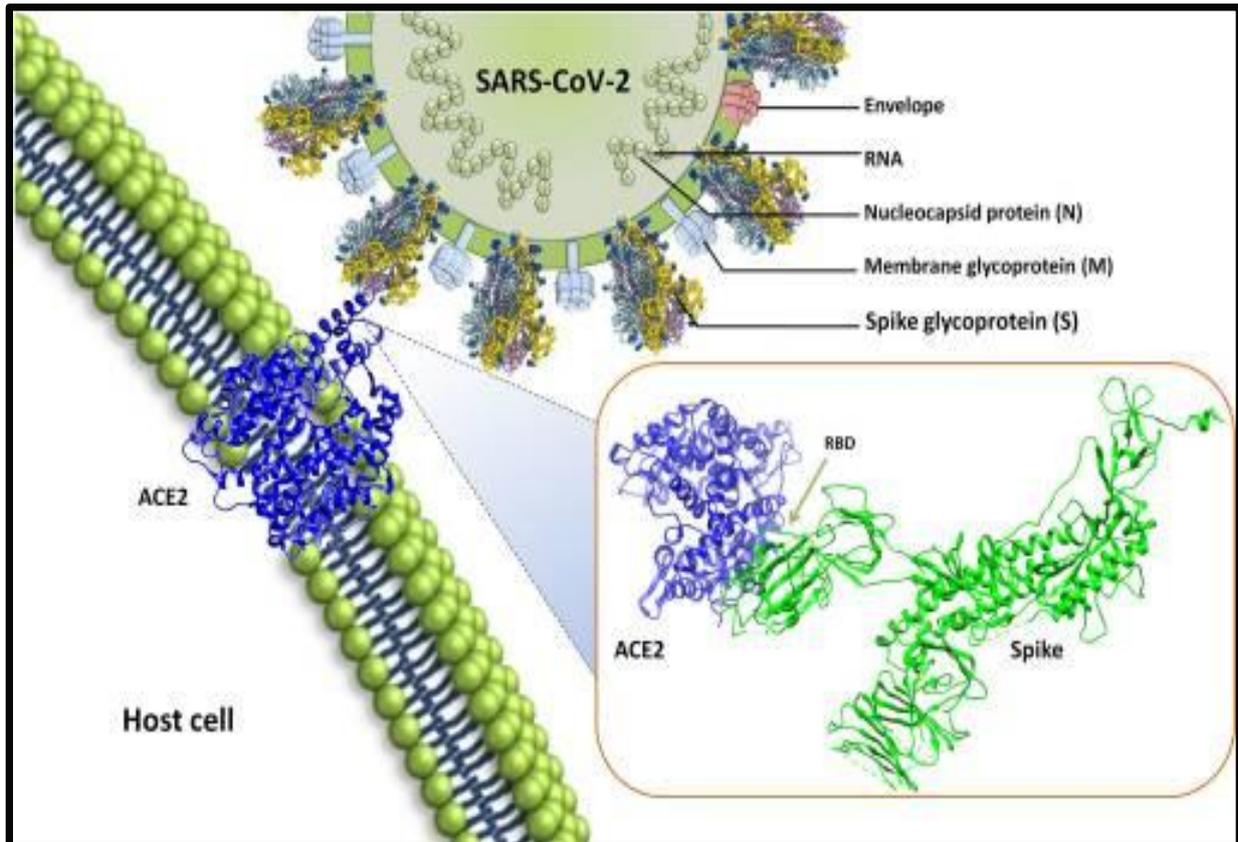


Figure 23 : Schéma illustrant la liaison entre le SARS-CoV-2 et le récepteur ACE2 présent à la surface des cellules cibles du coronavirus (cellule de l'hôte) (56).

Structurellement la protéine S est une très grande protéine transmembranaire qui s'assemble en trimères pour former des pointes de surface (SPIKE) distinctives des coronavirus (Figure 25) (15). Elle est divisée en deux parties fonctionnellement distinctes ; sous unité S1 et S2. La sous unité S1 contient un domaine de liaison au récepteur **RBD** « Receptor Binding Domain » qui est responsable de la reconnaissance et de la liaison avec le récepteur de surface cellulaire (ACE2). La sous-unité S2 est la "tige" de la structure qui assure la fusion des membranes virales et cellulaires lors de réarrangements structurels importants(57,58) (Figure 24).

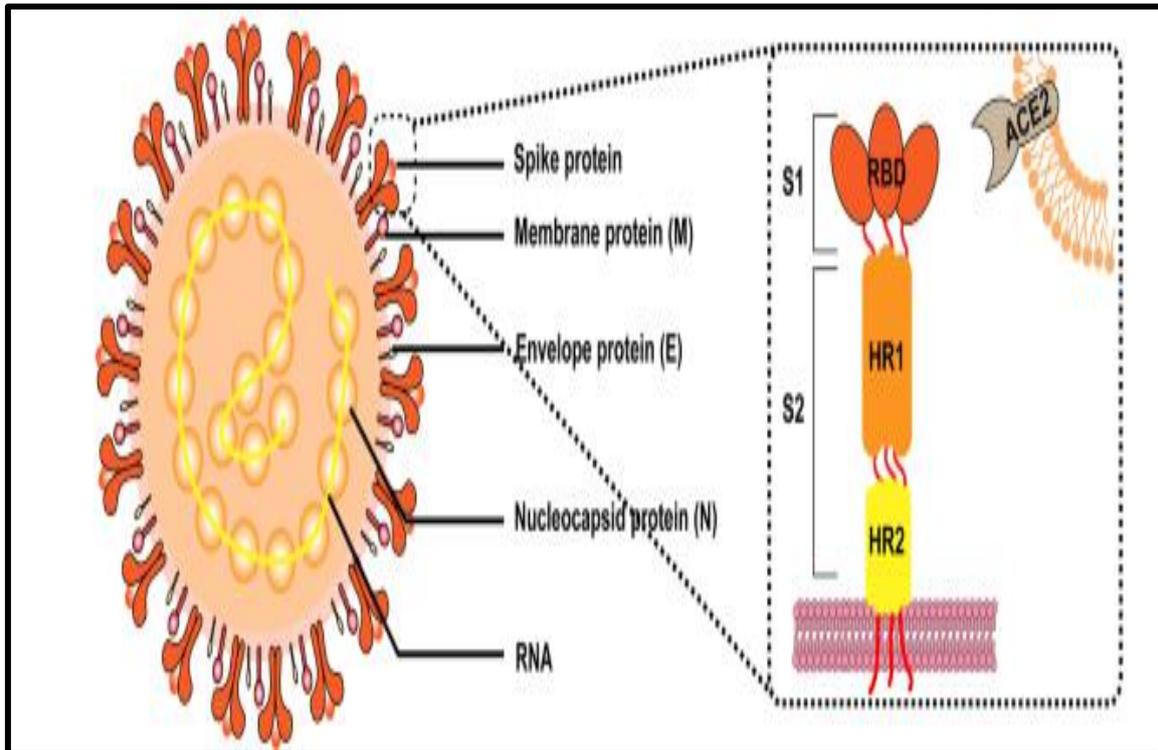


Figure 24 : Structure de la protéine S du SARS-COV-2 (59)

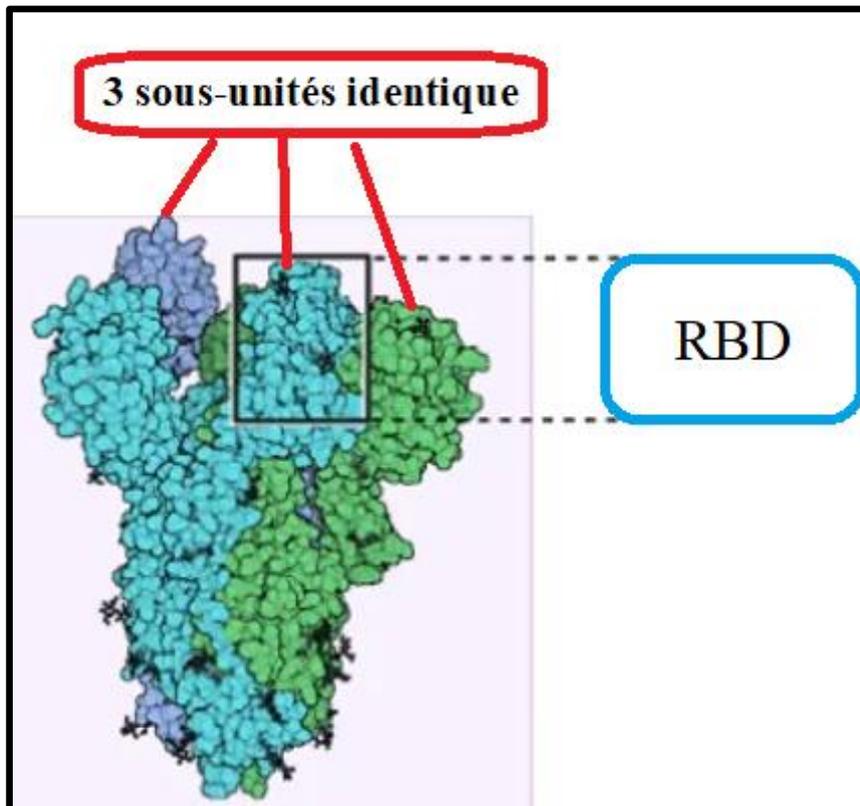
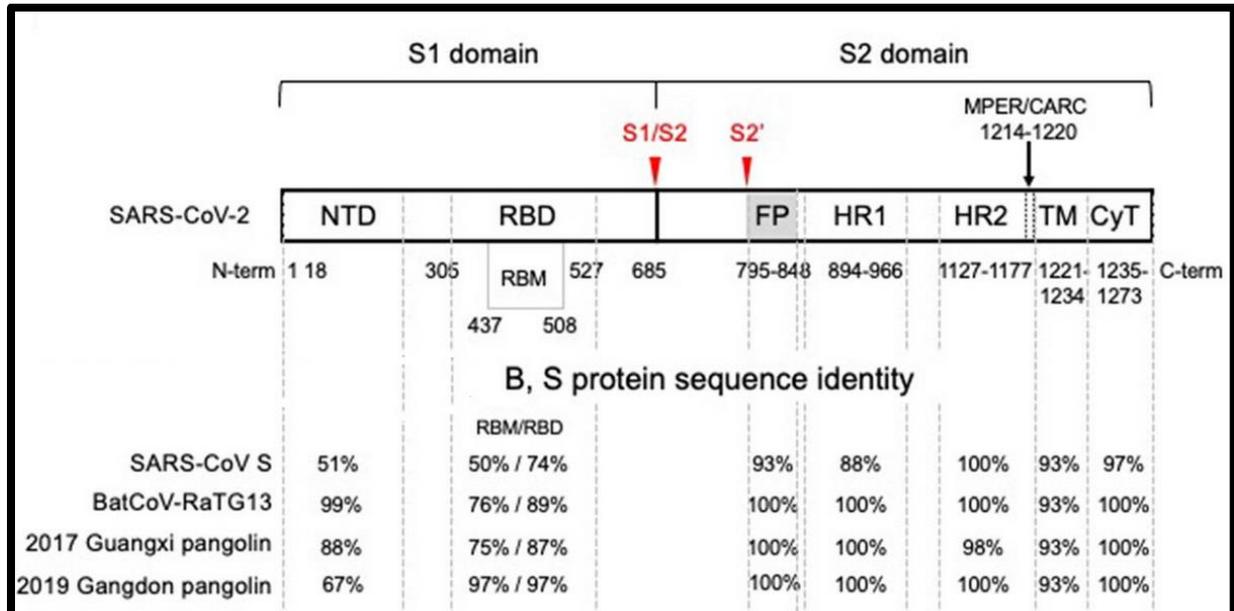


Figure 25 : Structure en trimère de la protéine S (60).

Le domaine S1 est la région la plus variable de la protéine S, tant entre les trois groupes de coronavirus qu'à l'intérieur de ceux-ci. Même parmi les souches d'une seule espèce de coronavirus, la séquence de la région S1 peut varier considérablement principalement au niveau de la région ; **RBD**. En revanche, la sous unité S2 est la région la plus conservée dans les trois groupes de coronavirus (Figure 26) (15).



NTD: N-terminal domain RBD: receptor binding domain RBM: receptor binding motif

FP: fusion peptide HR: heptad repeat TM: domaine transmembranaire

Figure 26 : similitude de la structure génomique de la protéine S des coronavirus (34).

3.3.2 La protéine nucléocapside « N »

Les nucléocapsides N des Bêta-coronavirus partagent une structure de domaine globale commune. C'est une phosphoprotéine de 43-50 kDa qui s'associe à l'ARN génomique pour former la nucléocapside hélicoïdale (11). Elle participe aussi à la réplication de l'ARN, à la formation des virions et à l'évasion immunitaire, elle interagit également avec les protéines PNS3 et la protéine M (Figure 27) (52).

3.3.3 La protéine membrane « M »

La protéine M est l'une des protéines les plus abondantes et la mieux conservée dans la structure du SARS-CoV-2, elle possède un triple faisceau en hélice et forme un seul domaine transmembranaire. La protéine M favorise l'assemblage et le bourgeonnement des particules

virales par l'interaction avec les protéines **N** et les protéines accessoires **3a** et **7a** (61) (Figure 27).

3.3.4 La protéine enveloppe « E »

La protéine d'enveloppe E est la principale composante de l'enveloppe du virus. C'est une protéine membranaire intégrale composée d'acides aminés de nature très hydrophobe (62), c'est la plus petite composante de la structure du SARS-CoV-2 (8 à 12 kDa) et est la moins abondante des protéines de l'enveloppe virale mais possède un rôle très important qui est de faciliter la production, la maturation, l'assemblage de l'enveloppe virale et la libération des virions (Figure 27) (11,52).

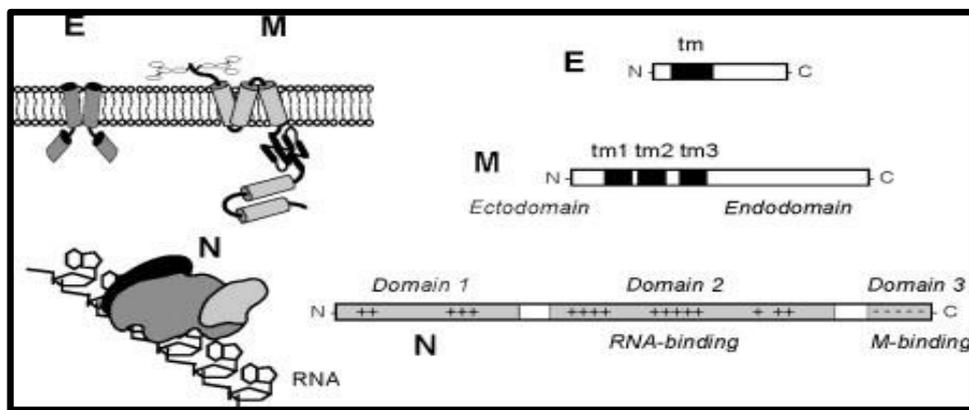


Figure 27: structure des protéines E, M, N (15).

3.3.5 Les protéines non structurales

Les protéines non structurales des coronavirus sont situées dans les deux premiers tiers du génome qui est composé de deux ORF (**ORF1a** et **ORF1b**). Le premier ORF qui contient environ 65 % du génome viral et le deuxième ORF sont traduits en deux polyprotéines **pp1a** et **pp1ab** qui sont elles-mêmes clivées par deux protéases virales permettant la libération des 16 protéines non structurales (PNS) (Figure 28)(63).

La PNS1 semble jouer un rôle dans la protection du virus contre le système immunitaire par la neutralisation de la réponse immunitaire dépendante de l'interféron. La PNS3 a un rôle majeur lors de la synthèse de l'ARN des coronavirus. La PNS5 est une protéase virale de type sérine qui permet la libération des PNS4 à 16 (57,63).

La PNS12 est l'enzyme clé lors de la réplication virale « l'ARN polymérase ARN-dépendante (**RdRp**) ». Il semble que **PNS8** en association avec **PNS7** ont également une activité ARN polymérase ARN-dépendante (63).

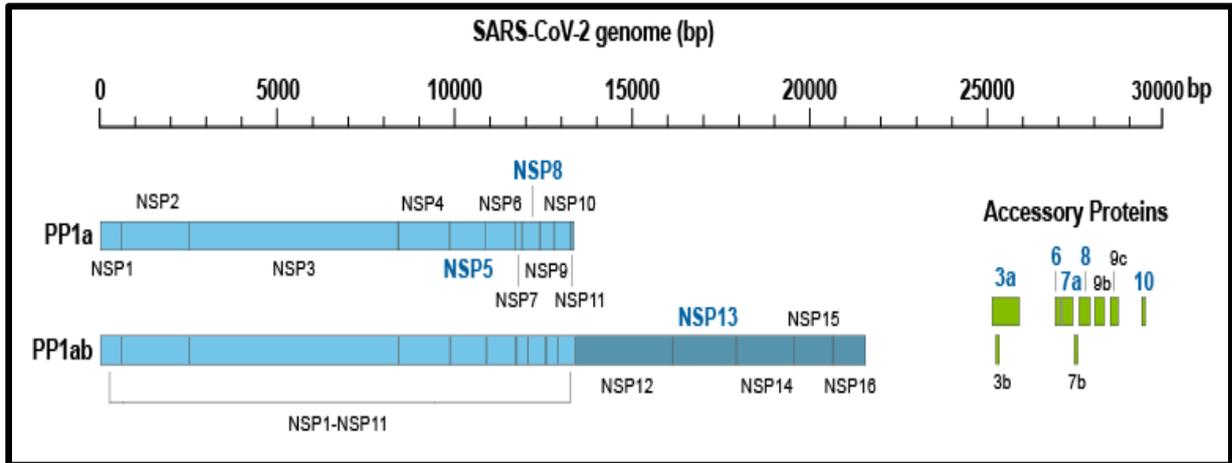


Figure 28 : les protéines non structurales des coronavirus (64).

3.4 Cycle de vie du SARS-COV-2

Comme pour d'autres coronavirus, le cycle de réplication du virus SARS-CoV-2 dans la cellule hôte peut être divisé en plusieurs étapes clés (65): (Figure 29).

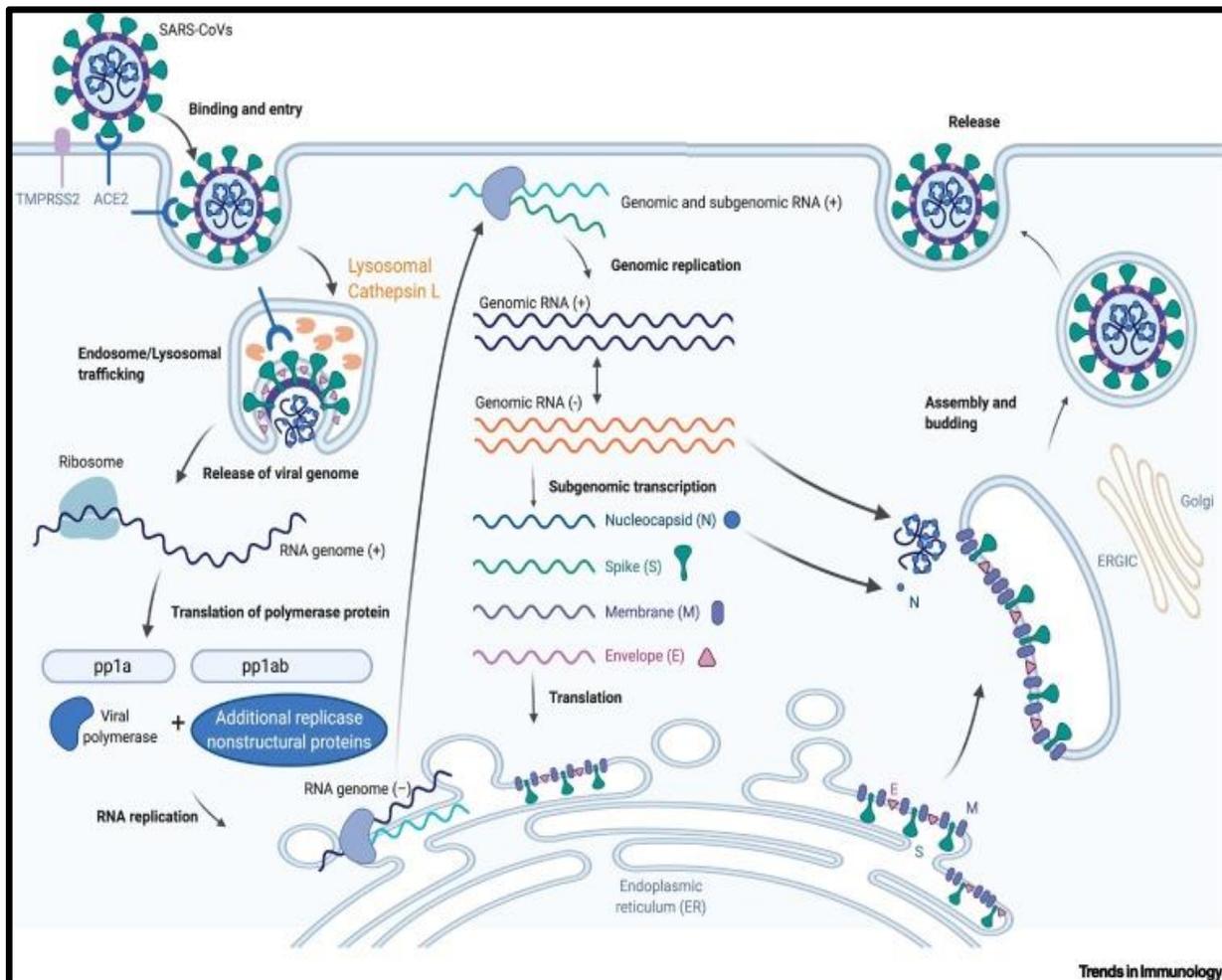


Figure 29 : Cycle de vie et réplication virale du SARS-COV-2 dans la cellule hôte (65).

3.4.1 L'attachement et l'entrée dans la cellule

Le SARS-CoV-2 infecte les cellules humaines par liaison de sa protéine S à la protéine transmembranaire cellulaire « les récepteurs ACE-2 », par l'intermédiaire du **RBD** « domaine de liaison au récepteur ». La liaison de la protéines S à l'ACE-2 entraîne un clivage protéolytique qui expose le peptide de fusion, ce clivage est effectué par la protéine cellulaire **TMPRSS-2** « **protéase transmembranaire sérine 2** » qui est nécessaire à l'amorçage de la protéine S aux récepteurs ACE-2 (16,66)(Figure 30).

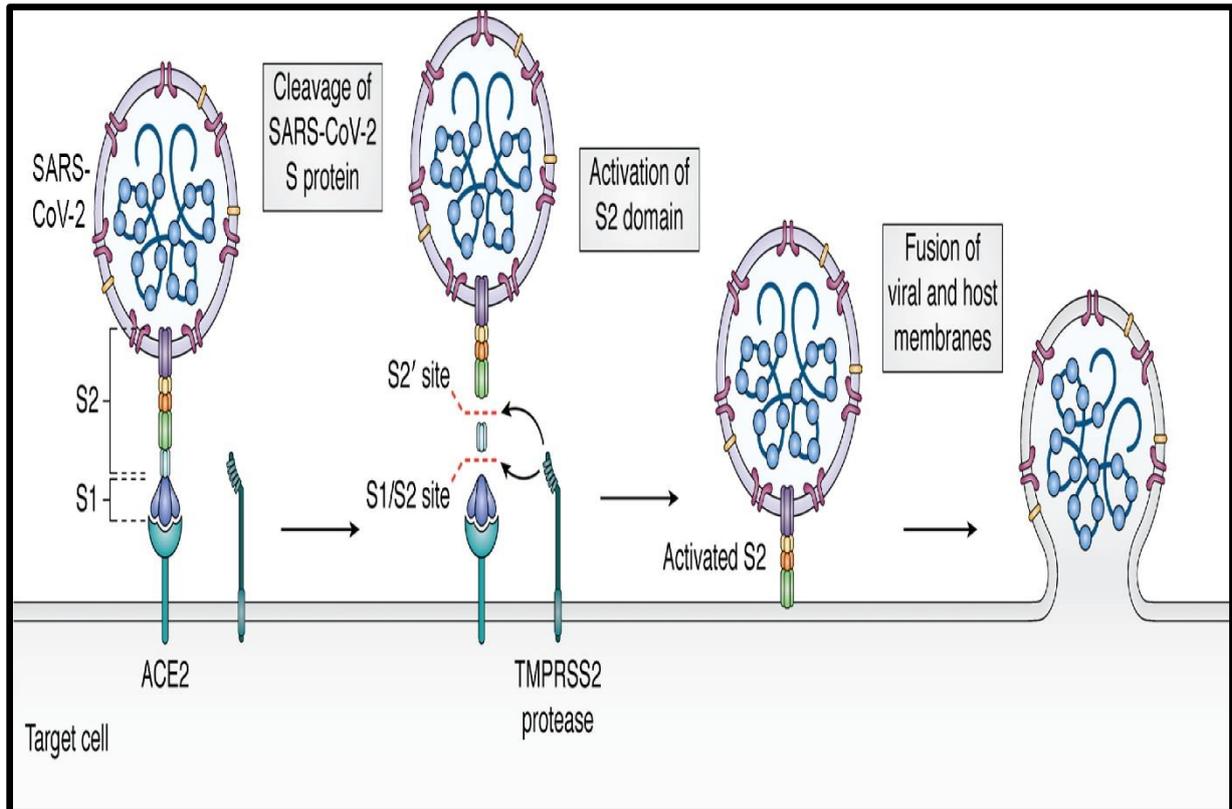


Figure 30 : Mécanisme de l'entrée du SARS-COV-2 à l'intérieur de la cellule hôte (48).

3.4.2 Traduction de l'ARN viral

Une fois l'entrée et la perte de l'enveloppe virale (décapsidation) par fusion avec la membrane cellulaire, l'ARNg positif (génomique) alors libéré dans le cytoplasme de la cellule hôte est traduit au niveau du ribosome par décalage du cadre de lecture « **Décalage Ribosomique** »(16). La traduction des régions ORF1a et ORF1b de l'ARN génomique produit respectivement deux polyprotéines, **pp1a** et **pp1ab** (Figure 31). Le clivage auto-protéolytique de ces deux polyprotéines génère 15 à 16 protéines non structurales qui possèdent des fonctions spécifiques(57) ; PNS12 code pour l'activité de l'ARN réplicase RdRp (ARN polymérase ARN dépendante) et en s'associant avec les deux cofacteurs PNS7 et PNS8 vont synthétiser des ARN

qui est le premier évènement de synthèse macromoléculaire après l'entrée du génome viral dans le cytoplasme. PNS3, 4 et 6 induisent le réarrangement de la membrane cellulaire pour former des vésicules à double membrane où est assemblé le complexe de transcription de la réplication du coronavirus (16,48,57).

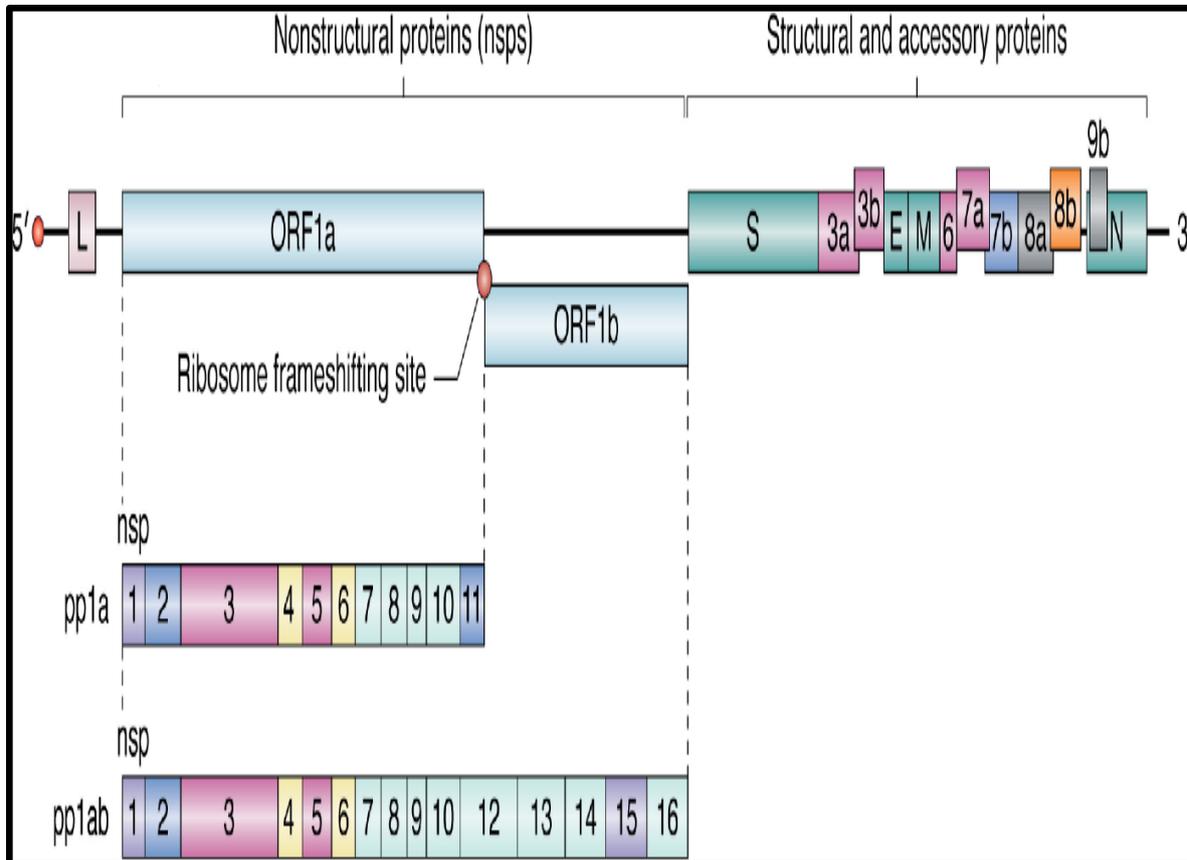


Figure 31 : organisation génomique et transcription virale du SARS-COV-2 (48).

3.4.3 Transcription et réplication génomique « RNA synthesis »

La réplication génomique virale est initiée par la transcription du brin positif de l'ARN virale en un brin d'ARN négatif qui vont servir de matrice pour la production de nouveaux ARN génomiques brins positifs (16). Les ARN positifs nouvellement synthétisés vont être soit traduits afin de générer plus de PNS, soit sont assemblés en nouveaux virions.(57)(Figure 32,33). Le complexe RdRp réalisera également la transcription de l'ensemble des domaines ORF en ARN messagers (ARNm) dits « sous-génomiques » qui par la suite vont être traduits par les ribosomes en protéines structurales et accessoires (11).

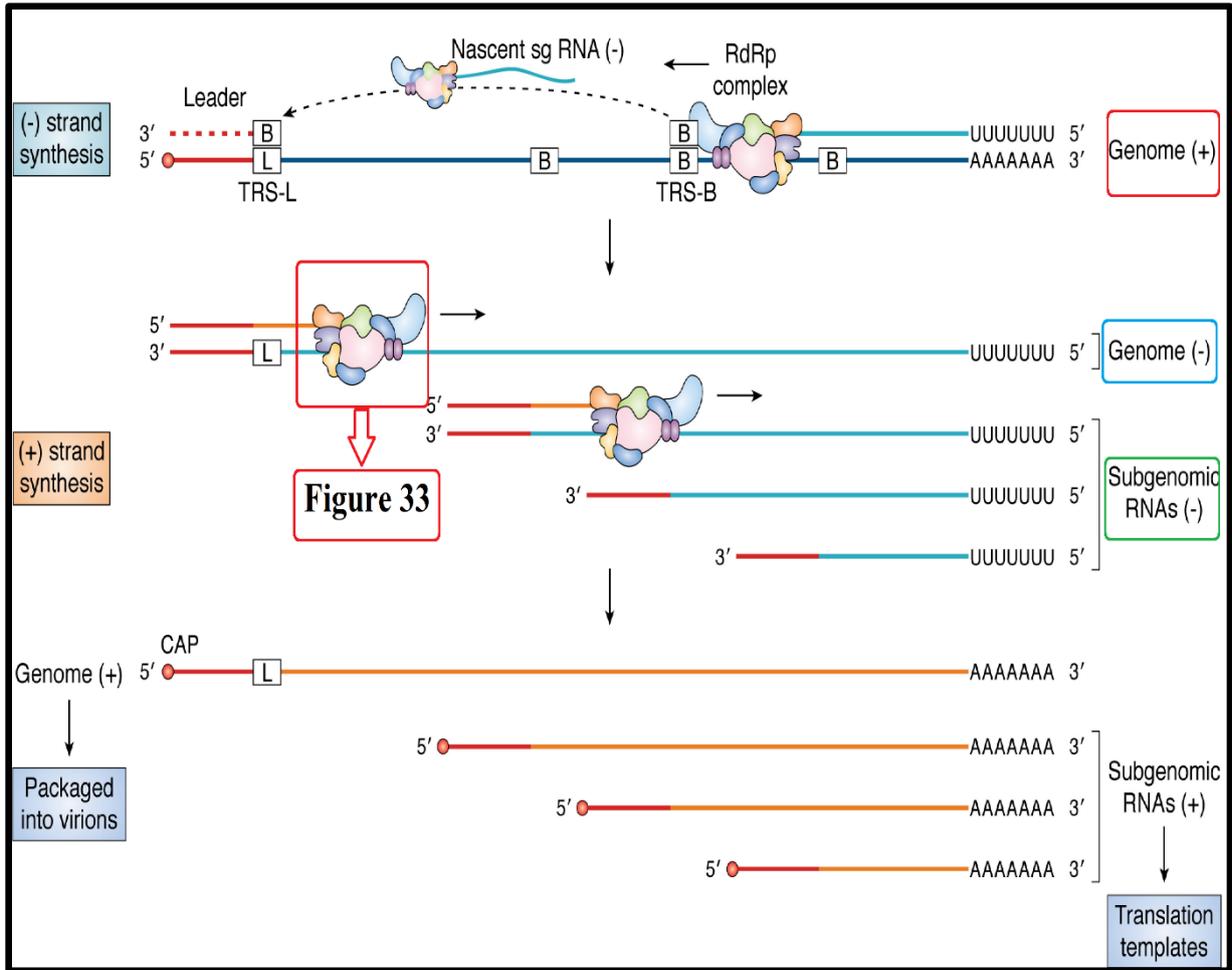


Figure 32 : schéma détaillant les différentes étapes de synthèse de l'ARN viral (48).

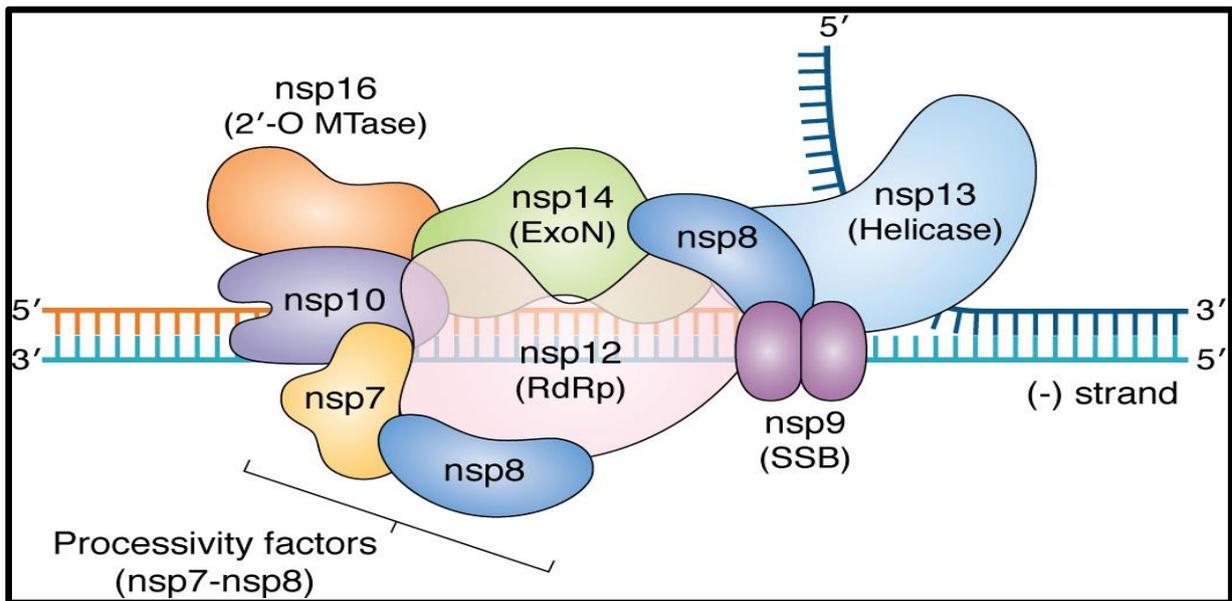


Figure 33 : Structure du complexe RdRp. « ARN polymérase ARN dépendante »(48).

3.4.4 Traduction des protéines structurales, assemblage et libération des virions

Les protéines structurales une fois synthétisées vont s'assembler et participer à la formation de nouveaux virions soit au niveau du réticulum endoplasmique (c'est le cas des protéines S, E, M et HE), soit au niveau de l'appareil de Golgi (protéine N), ces nouveaux virions vont sortir de la cellule infectée par exocytose et vont infecter d'autres cellules (invasion virale) (15,16,57).

3.5 Les variants du SARS-COV-2

Comme les autres virus à ARN le SARS-CoV-2 est sujet à des évolutions génétiques qui peuvent donner naissance à de multiples variants pouvant présenter des caractéristiques différentes de celles de la souche initiale. Certaines modifications peuvent être sans conséquence alors que d'autres peuvent être avantageuses. Les mutations se produisent sur toute la longueur du génome du virus bien qu'environ 50 % se produisent au niveau du gène de la protéine Spike (gène S) (33).

Les mutations définissent chaque nouvelle lignée du SARS-COV-2. Lorsque les mutations se produisent dans la protéine S et en particulier dans la région RBD, elles peuvent affecter la liaison avec les récepteurs de l'hôte ou avec les anticorps (33).

Avec l'émergence rapide de plusieurs variants, l'OMS a établi un système de classification pour distinguer les variants émergents du SARS-CoV-2 :

3.5.1 Le variant G614

En ce moment cette mutation se trouve dans presque toutes les séquences du SARS-COV-2 dans le monde, c'est la première mutation observée dans le SARS-COV-2 au niveau du gène de la protéine S qui consiste en une substitution de l'acide aspartique par la glycine en position 614 améliorant la liaison entre la protéine S et le récepteur ACE2 des cellules cibles.(Figure 34) (33,67).

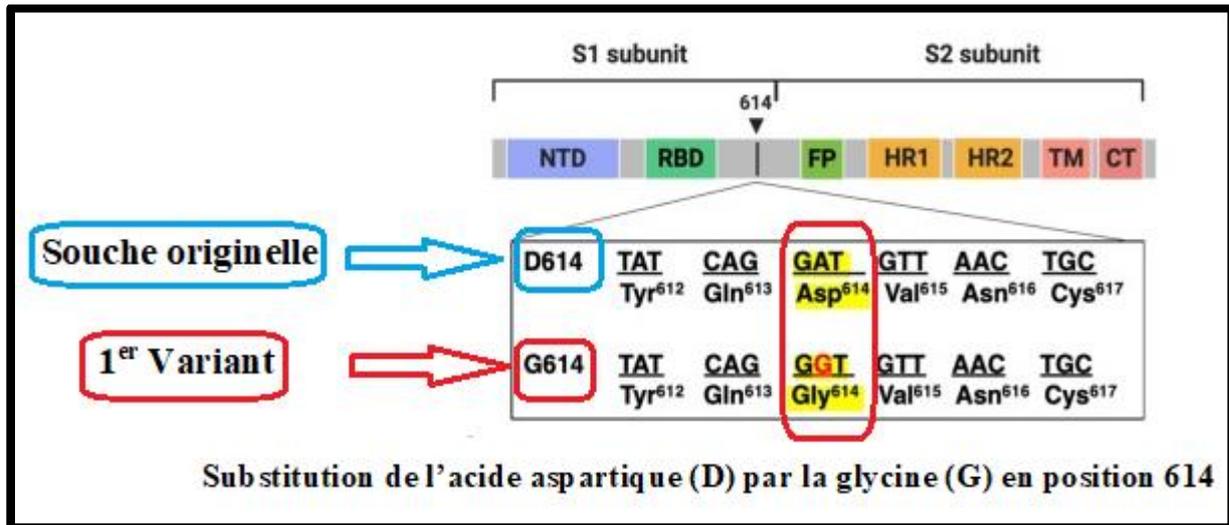


Figure 34 : La mutation D614G au niveau de la protéine S du SARS-COV-2 (68).

3.5.2 Les variants préoccupants « variant of concern »

Ce sont les variants pour lesquelles il existe des preuves de changements phénotypiques par rapport à la souche initiale et qui entraînent (33):

- Une augmentation de la transmissibilité du virus (36).
- Une réduction significative de la neutralisation du virus par les anticorps générés lors d'infection ou vaccination antérieure (33).
- Une diminution de l'efficacité des traitements ou des vaccins (36).
- Une augmentation de la sévérité de la maladie (augmentation des hospitalisations ou des décès) (33).
- L'échec de détection par des tests spécifiques (36).

3.5.2.1 Le variant Anglais « Alpha »

Le variant Alpha a été identifié pour la première fois en décembre 2020 au Royaume-Uni dans un test commercial fréquemment utilisé, le « TaqPath RT-PCR », qui permet par la technique RT-PCR « Reverse transcriptase PCR- polymerase chain reaction » de détecter trois cibles différentes au niveau de l'ARN viral, cependant le variant alpha comporte la délétion 69-70 ou encore désignée sous le nom Δ H69/V70 car elle provoque la perte des acides aminés histidine en position 69 et valine en position 70 au niveau de la protéine S entraînant ainsi l'obtention de deux signaux au lieu de trois et donc l'absence de détection du gène S (S gene target failure ou SGTF) (19).

Le variant alpha comprend dix-sept mutations dans le génome dont neuf mutations se trouvent au niveau du gène de la protéine S (19) (Figure 35) :

- **S : Délétion Δ H69-V70**
- **S : Délétion Δ Y144** : Délétion de la tyrosine en position 144. Elle confère une résistance aux anticorps monoclonaux 4A8.
- **S : N501Y** : signifie qu'en position 501 de la chaîne des acides aminés de la protéine S une tyrosine (Y) a remplacé une asparagine (N).
- **S : A570D** ; substitution de l'alanine par l'acide aspartique en position 570.
- **S : D614G**.
- **S : P681H** : substitution de la proline par l'histidine en position 687. Mutation proche du site de clivage S1/S2.
- **S : T716I** : substitution de la thréonine par l'isoleucine en position 716.
- **S : S982A** : substitution de sérine par l'alanine en position 982.
- **S : D1118H** : substitution de l'acide aspartique par l'histidine (H) en position 1118 (36).

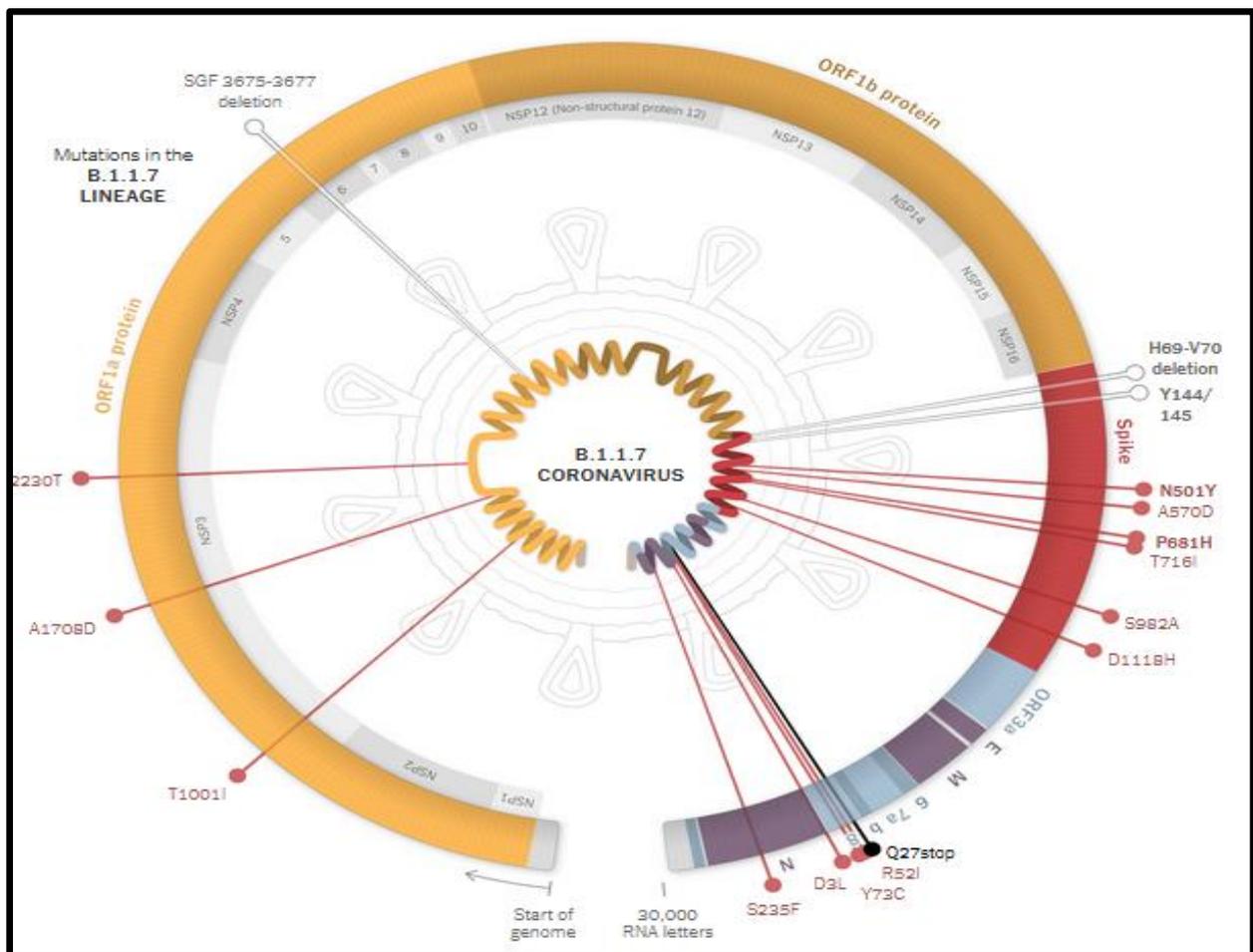


Figure 35 : les mutations clés présentes dans le génome du variant Anglais (30).

Etant donné qu'il faut trois protéines S pour former une pointe donc chaque mutation apparaît à trois endroits (Figure 36) (30).

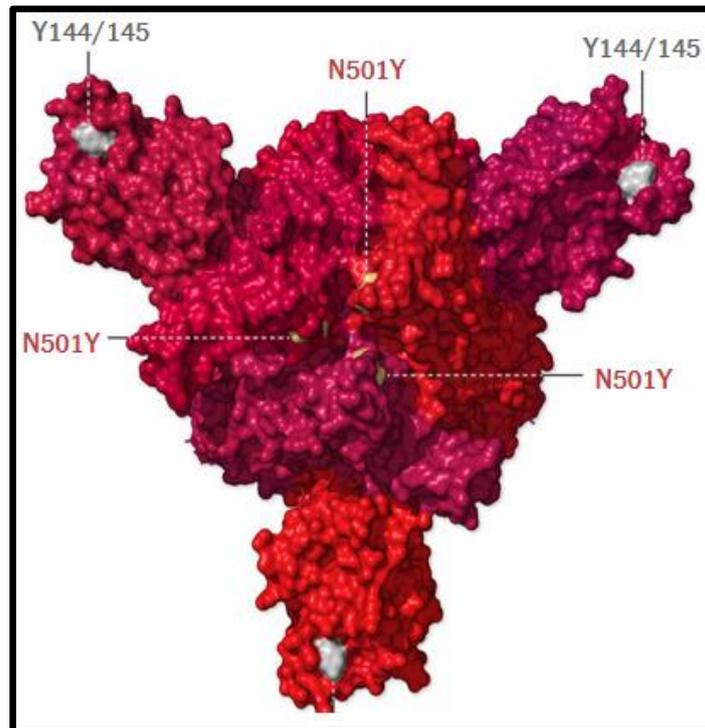


Figure 36 : les mutations de la protéine S du variant Anglais (30).

3.5.2.2 Le variant Sud-Africain « Béta » :

Le variant Béta comprend neuf mutations (L18F, D80A, D215G, R246I, K417N, E484K, N501Y, D614G et A701V) dans la protéine S dont trois mutations sont situées dans le RBD (K417N, E484K et N501Y) (30) :

- **S : L18F*** : substitution de la leucine (L) par la phénylalanine (F) en position 18
- **S : D80A** : substitution de l'acide aspartique (D) par l'alanine (A) en position 80.
- **S : D215G** : substitution de l'acide aspartique (D) par la glycine (G) en position 215
- **S : Délétion L241/L242** : Délétion de la leucine (L) en position 241/242.
- **S : Délétion A243** : Délétion de l'alanine (A) en position 243.
- **S : K417N** : substitution de la lysine (K) par l'asparagine (N) en position 417.
- **S : E484K** : substitution de l'acide glutamique (E) par la lysine (K) en position 484.
- **S : N501Y, D614G**
- **S : A701V** : substitution de l'alanine (A) par la valine en position 701(36).

(*): mutation présente que dans certaines sous souche.

3.5.2.3 Variant Brésilien « Gamma »

Le variant Gamma héberge douze mutations dans la protéine S (L18F, T20N, P26S, D138Y, R190S, H655Y, T1027I, V1176F, K417T, E484K et N501Y) dont trois mutations (L18F, K417N, E484K) sont localisées dans le RBD similaire au variant Béta (36)

- S : **L18F**
- S : **T20N** : substitution de la thréonine (T) par l'asparagine (N) en position 20.
- S : **P26S** : substitution de la proline (P) par la sérine (S) en position 26.
- S : **D138Y** : substitution de l'acide aspartique (D) par la tyrosine (Y) en position 138.
- S : **R190S** : substitution de l'arginine (R) par la sérine en position 190.
- S : **K417T, E484K, N501Y, D614G**
- S : **H655Y** : substitution de l'histidine par la tyrosine en position 655.
- S : **T1027I** : substitution de la thréonine (T) par l'isoleucine (I) en position 1027.
- S : **V1176F** : substitution de la valine (V) par la phénylalanine (F) en position 1176(36).

Le variant Brésilien semble présenter un profil de mutations similaire au variant sud-africain (K417N/T, E484K, N501Y) (69) (Figure 37)

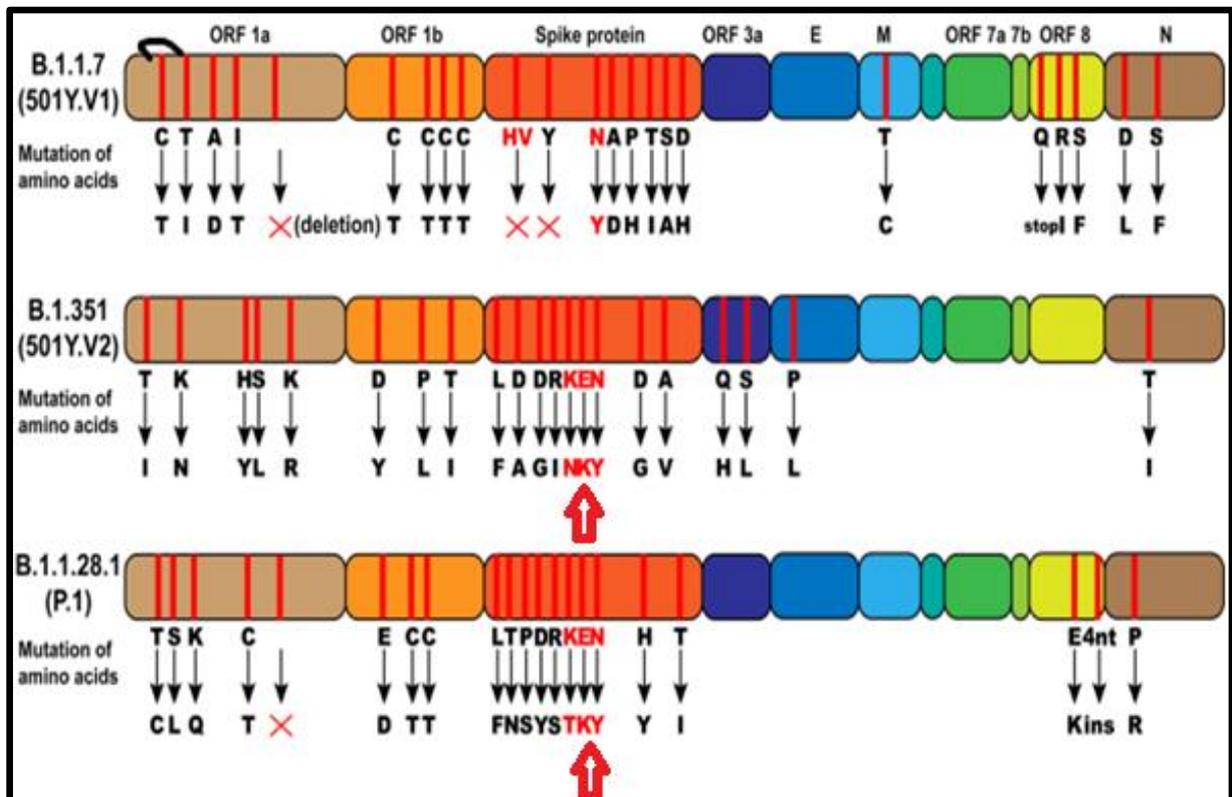


Figure 37 : les mutations observées dans les variants « Alpha », « Béta », « Gamma » (59).

3.5.2.4 Le variant Indien « Delta »

Il a été détecté pour la première fois en Inde en décembre 2020 et est porteur de plus d'une douzaine de mutations dont 6 mutation au niveau de la protéine S (30) :

- **S : E154K** : substitution de l'acide glutamique (E) par la lysine (L) en position 154.
- **S : L452R** : Substitution de la leucine (L) par l'arginine (R) en position 452.
- **T478K** : Substitution de la thréonine (T) par la lysine (K) en position 478
- **S : E484Q** : substitution de l'acide glutamique (E) par la glutamine.
- **S : D614G**
- **S : P681R** : substitution de la proline par l'arginine en position 681,
- **S : Q1071H** : substitution de la glutamine par l'histidine en position 1071 (36).

3.5.3 Les variants à suivre (variant of interest)

Ce sont des variants présentant des marqueurs génétiques spécifiques associées à des changements phénotypiques confirmés ou suspectés qui peuvent entraîner :

- Des modifications au niveau des récepteurs de liaison.
- Une neutralisation moindre par les anticorps générés lors d'infection naturelle ou vaccination antérieure.
- Une efficacité réduite des traitements.
- Une augmentation de la transmissibilité ou de la gravité de la maladie (36,69,70).

3.5.3.1 Les variants Américains

3.5.3.1.1 Le variant Epsilon

Ces deux variants sont apparus pour la première fois aux Etats-Unis en juin 2020 et sont passées de 0 % à plus de 50 % des cas recensés entre le 1er septembre 2020 et le 29 janvier 2021. Actuellement Ils sont considérés comme variants préoccupants aux Etats-Unis alors que l'OMS les classent comme variants à suivre (69).

❖ Le B.1.427 comprend deux mutations :

- **D614G, L452R**

❖ Le B.1.429 comprend quatre mutations :

- **S13I** : substitution de la sérine (S) par l'isoleucine en position 13.
- **W152C** : substitution du tryptophane (W) par la cystéine en position 152.
- **L452R, D614G** (36).

3.5.3.1.2 Le variant Iota « B.1.526 »

Détecté pour la première fois à New-York en novembre 2020, il se présente sous deux sous souches : la première comprend la mutation E484K dans la protéine S. La deuxième porte la mutation S477N au niveau de la protéine S

- **S : L5F*** : substitution de la leucine (L) par la phénylalanine (F) en position 5.
- **S : T95I** : substitution de la thréonine (T) par l'isoleucine (I) en position 95.
- **S : D253G** : substitution de l'acide aspartique (D) par la glycine (G) en position 253.
- **S : S477N*** : substitution de sérine (S) par l'asparagine (N) en position 477.
- **S : E484K*, D614G**
- **S : A701V*** : substitution de l'alanine (A) par la valine (V) en position 701(36).

3.5.3.2 Le variant Nigérien « Eta »

Initialement détecté au Nigéria et aux Etats-Unis en décembre 2020, Il est porteur de la mutation E484K et de deux délétions (Δ H69/V70, Δ 144) dans la protéine Spike similaire à celles retrouvées chez le variant Alpha. Il comprend les mutations suivantes (19):

- **S : Q52R** : substitution de glutamine (G) par l'arginine (R) en position 52.
- **S : A67V** : substitution de l'alanine (A) par la valine (V) en position 67.
- **S : E484K**
- **S : Délétion H69/V70**
- **S : Délétion Y144, D614G**
- **S : Q677H** : substitution de la glutamine (Q) par l'histidine (H) en position 677.
- **S : F888L** : substitution de phénylalanine (F) par la leucine (L) en position 888(36).

3.5.4 Les variants en cours d'évaluation (variant under monitoring)

Ce sont les variants pour lesquelles il y a absence d'éléments virologiques, épidémiologiques ou cliniques indiquant un impact sur la santé publique (augmentation transmission, sévérité des symptômes) malgré la présence de mutations partagées avec un ou plusieurs variants préoccupants ou variants à suivre. Actuellement trois variants figurent sur la liste des variants en cours d'évaluation (36).

3.5.4.1 Le variant B.1.214.2 :

Apparus dès avril 2021 en Belgique, il est caractérisé par une insertion de trois acides aminés et quatre mutations au niveau du gène de la protéine S, et la présence de plusieurs mutations et délétions dans les gènes codant pour d'autres protéines (délétion 3675-3677 dans le gène ORF1ab retrouvée chez les quatre VOC) (19).

3.5.4.2 Le variant N.9 :

Il aurait émergé au Brésil et est porteur de la mutation E484K ainsi que de plusieurs délétions dans la protéine S.

Les variants du SARS-COV-2 dérivent tous de la souche initiale (WT) isolée à Wuhan en chine et possèdent pour la plupart des mutations communes au niveau du génome (Figure 38)(33).

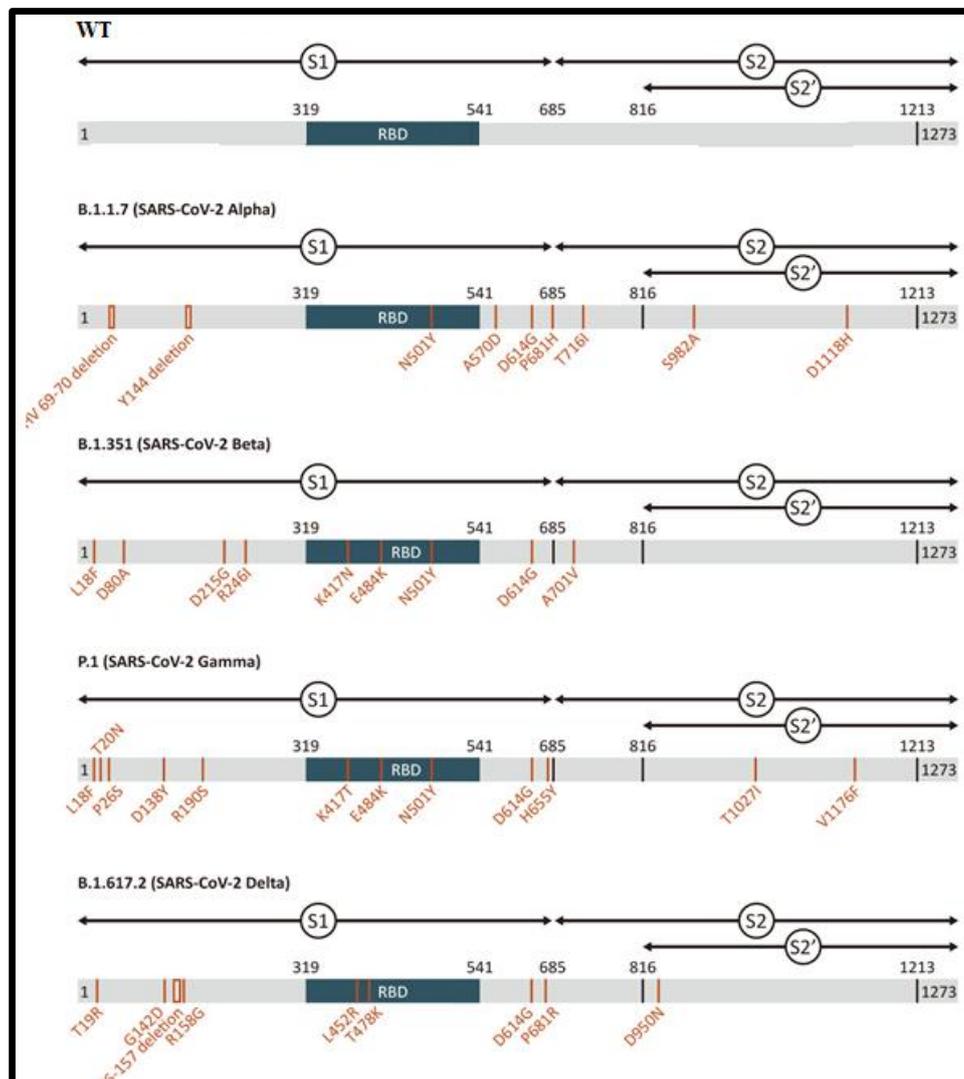


Figure 38 : Les principales mutations observées dans le SARS-COV-2.

3.6 Pathogénie du SARS-COV-2

3.6.1 Le système rénine-angiotensine

Le SRA (système rénine-angiotensine) joue un rôle important dans la régulation de la pression artérielle essentiellement par l'action de deux protéines ; ACE1 et ACE2 (71).

Dans les poumons, l'angiotensine I est convertie par l'ACE1 en angiotensine II qui se lie au récepteur AT1 (AT1R) provoquant ainsi une vasoconstriction, une hypertension et une inflammation. L'angiotensine I peut aussi être converti en angiotensine IV par le récepteur AT4 (AT4R) qui entraîne une thrombose. Le rôle de l'ACE2 est d'inactiver l'angiotensine II en la convertissant en angiotensine (1-7) qui se lie au récepteur MAS (Mas-related G protein-coupled receptors) et entraîne une vasodilatation et l'hypotension (71,72).

3.6.2 Système rénine angiotensine et SARS-COV-2

La liaison de la protéine S du SARS-COV-2 au récepteur ACE2 entraîne la diminution de la disponibilité de l'ACE2 ce qui résulte en une moindre dégradation de l'angiotensine II qui conduit aux anomalies observées dans l'étude SARS-COV-2, à savoir une lésion pulmonaire aiguë avec une vasoconstriction locale facilitant le Syndrome de détresse respiratoire aiguë, une lésion myocardique et une thrombose (Figure 39) (71,73).

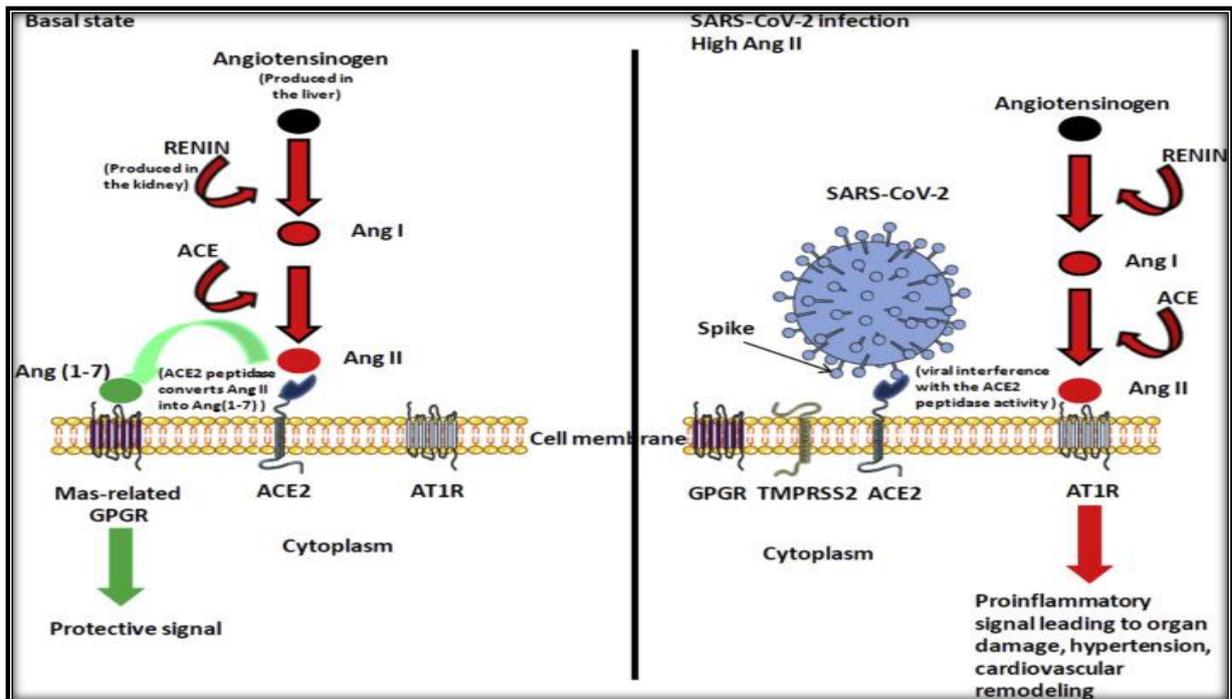


Figure 39 : Rôle du SRA dans la virulence du SARS-COV-2 (74).

3.6.3 Mécanisme d'échappement du SARS-COV-2 à l'immunité innée de l'hôte

la principale fonction de la structure virale est d'infecter les cellules et de prendre le contrôle de la machinerie de la cellule hôte pour se multiplier et produire de grandes quantités de virions qui n'auront qu'un seul but ; infecter les autres cellules de l'organisme et répéter le processus en faisant appels à plusieurs mécanismes de survies (50,75) :

- ❖ Échappement à la reconnaissance antigénique par les PRR (**Pattern Recognition Receptor** « récepteur de reconnaissance de pathogènes ») via la production de vésicules à double membrane abritant le complexe de réplication virale (76).
- ❖ Diminution de la signalisation des PRR soit par liaison compétitive de la protéine N à TRIM25 « **Tripartite Motif Containing 25** » bloquant ainsi la signalisation du RIG-1 « **Retinoic Acid-inducible Gene I** »(50,76), soit par l'inhibition du MDA-5 « **melanoma differentiation-associated protein 5** » par la PNS16 (76).
- ❖ Blocage de la signalisation des interférons : la PNS1 peut entraver la réponse interféron par le biais de plusieurs mécanismes tels que (50,76,77) :
 - L'inhibition de l'ARNi (ARN interférent)
 - La dégradation de l'ARNm de l'hôte
 - L'inhibition de la protéine STAT1 « **Signal Transducers and Activators of Transcription** ».
- ❖ Blocage de la signalisation du facteur nucléaire kappa B (NF-KB) par la PNS3 pour le SARS-COV-2, Protéines PLP (papain-like protease) pour le SARS-COV-1, ORF4b et ORF5 pour MERS-COV (Figure 40) (50,76,77).

Ainsi grâce à ces mécanismes complexes le génome ARN viral n'est pas reconnu par les cellules du système immunitaire (76).

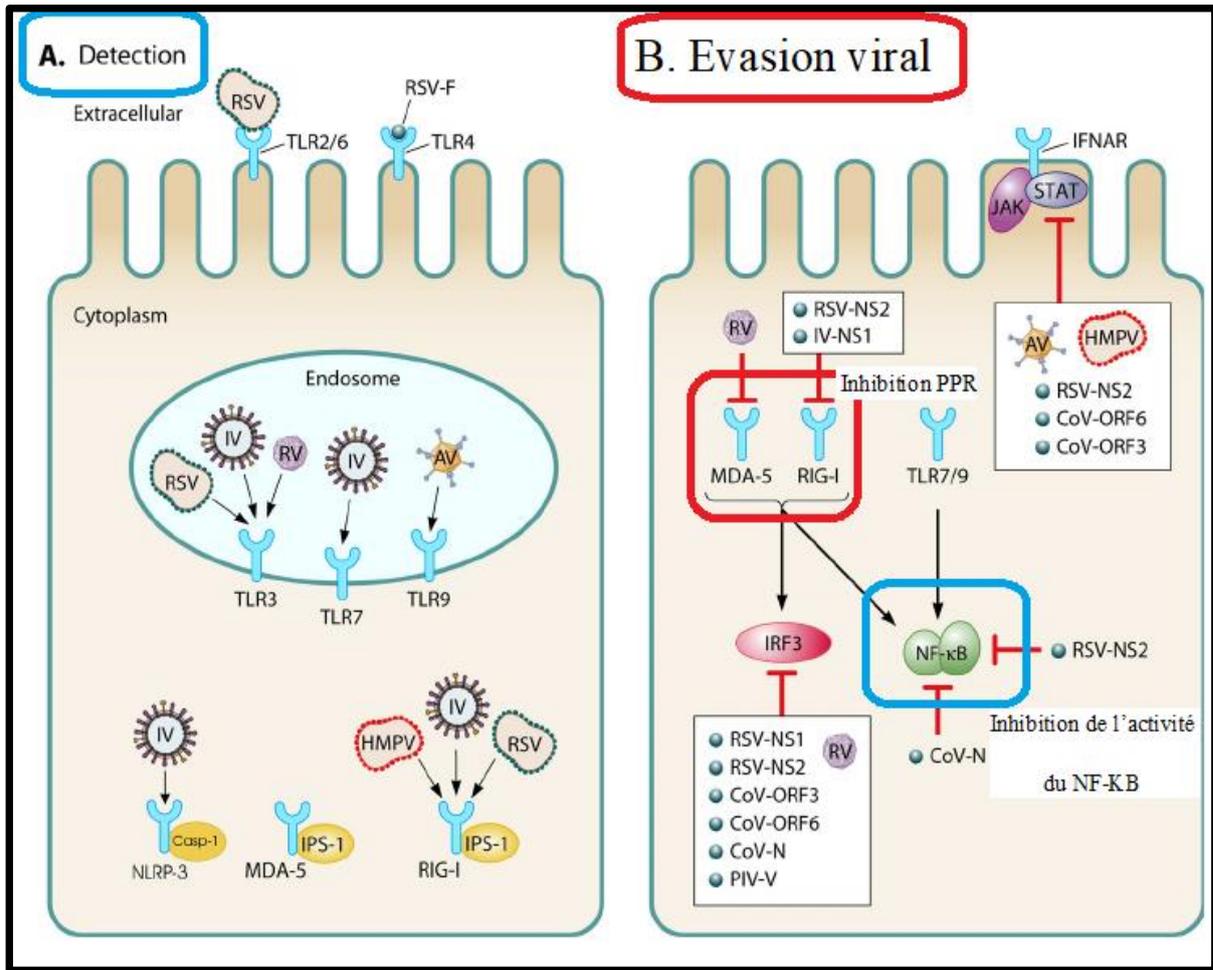


Figure 40 : mécanisme de la détection immunitaire et de l'évasion virale (78).

3.7 Symptômes du SARS-COV-2

Les symptômes de l'infection par le SARS-COV-2 apparaissent après une période d'incubation d'environ 5 jours (entre 2 et 14 jours)(79). L'un des principaux problèmes de la pandémie du SARS-COV-2 est la diversité des symptômes de la maladie qui peuvent varier d'une forme asymptomatique (80% des cas) à une pneumonie bilatérale sévère avec détresse respiratoire aiguë pouvant entraîner la mort (80) (Figure 41).

Les manifestations cliniques les plus courantes sont (79) :

- La fièvre
- Une toux sèche dans 67 % des cas et grasse dans 33% des cas.
- La fatigue
- Douleur musculaire et articulaire
- La dyspnée (essoufflement).
- Frissons ou état fébrile

D'autres symptômes moins fréquents sont : Les maux de tête, Les maux de gorge, l'hémoptysie (expectoration de sang), nausée, vomissement et diarrhée Conjonctivite (1).

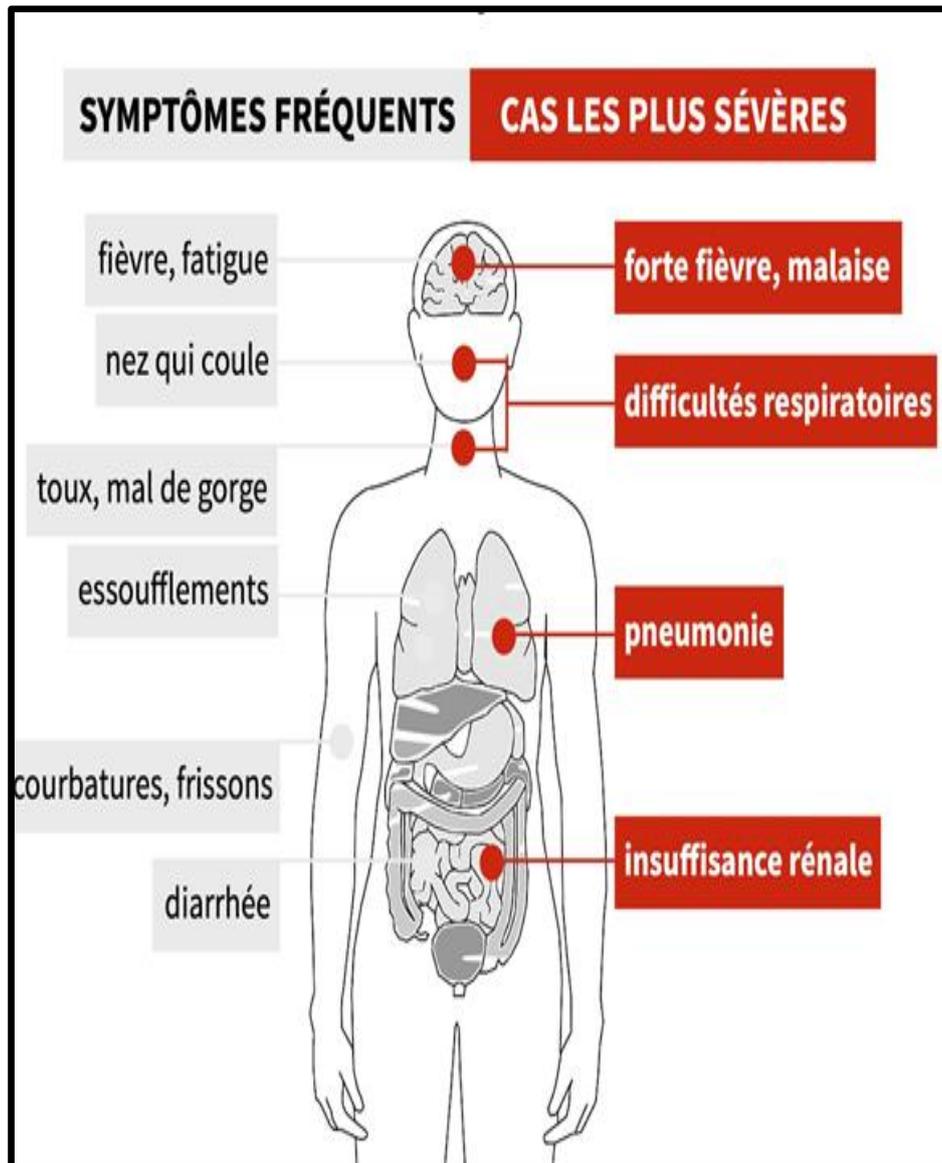


Figure 41 : les manifestations cliniques du SARS-COV-2 (Source OMS).

Bien que les symptômes du SARS-COV-2 soit dangereux, il n'entraîne pas une forte mortalité. Cependant la dangerosité du SARS-CV-2 réside dans la surmortalité qu'il entraîne pendant plusieurs mois après la guérison des patients. Ainsi les survivants ont 59% plus de risques de mourir dans les six mois suivant l'infection par rapport à une personne non infectée. Une étude a montré que 75 % des patients souffraient encore de différents symptômes six mois après l'infection et qu'il était possible de développer des complications à long terme potentiellement mortelle (Figure 42) (81).

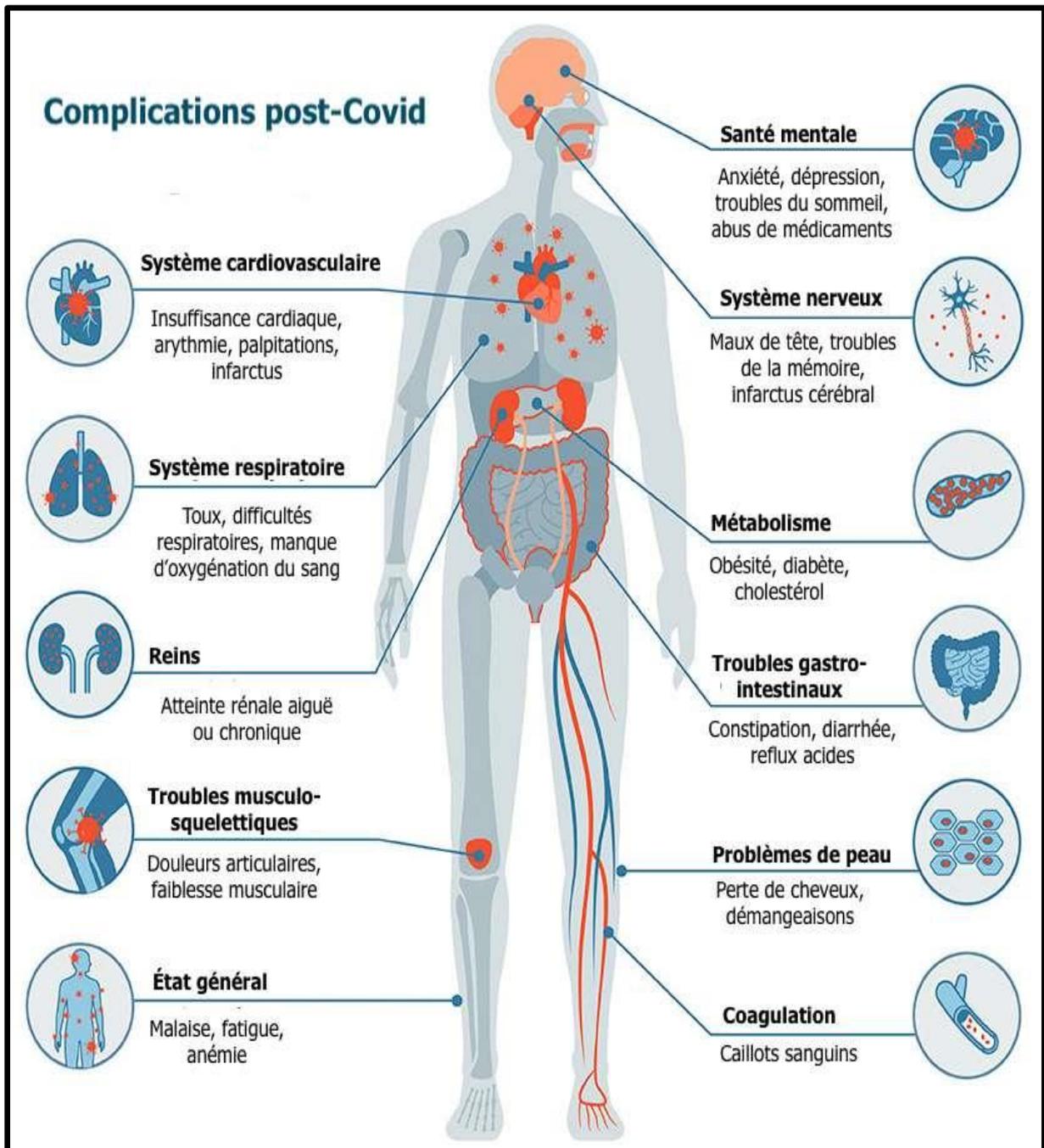


Figure 42 : Complications après l'infection au SARS-COV-2 (81).

3.8 Physiopathologie du SARS-COV-2

La principale pathogénie du SARS-COV-2 est une pneumonie sévère, une ARNémie, la présence d'un infiltrat en verre dépoli (image radiographique) et des lésions cardiaques aiguës(79). L'infection débute dans les voies respiratoires supérieures et progresse dans les cas graves vers les régions inférieures du poumon, les gouttelettes respiratoires porteuses du SARS-CoV-2 infectent les cellules épithéliales et endothéliales, les neurones et les macrophages

pulmonaires. La réplication virale et la libération des virions entraînent une pyroptose qui induit un dysfonctionnement de la réponse immunitaire innée (Figure 43) provoquant ainsi une libération excessive des cytokines « choc cytokinique », ces cytokines en excès entraînent une aggravation de l'inflammation, une augmentation de la vasodilatation, une augmentation de la perméabilité capillaire et de l'hypoxémie qui peut souvent conduire à la défaillance de plusieurs organes(16,79).

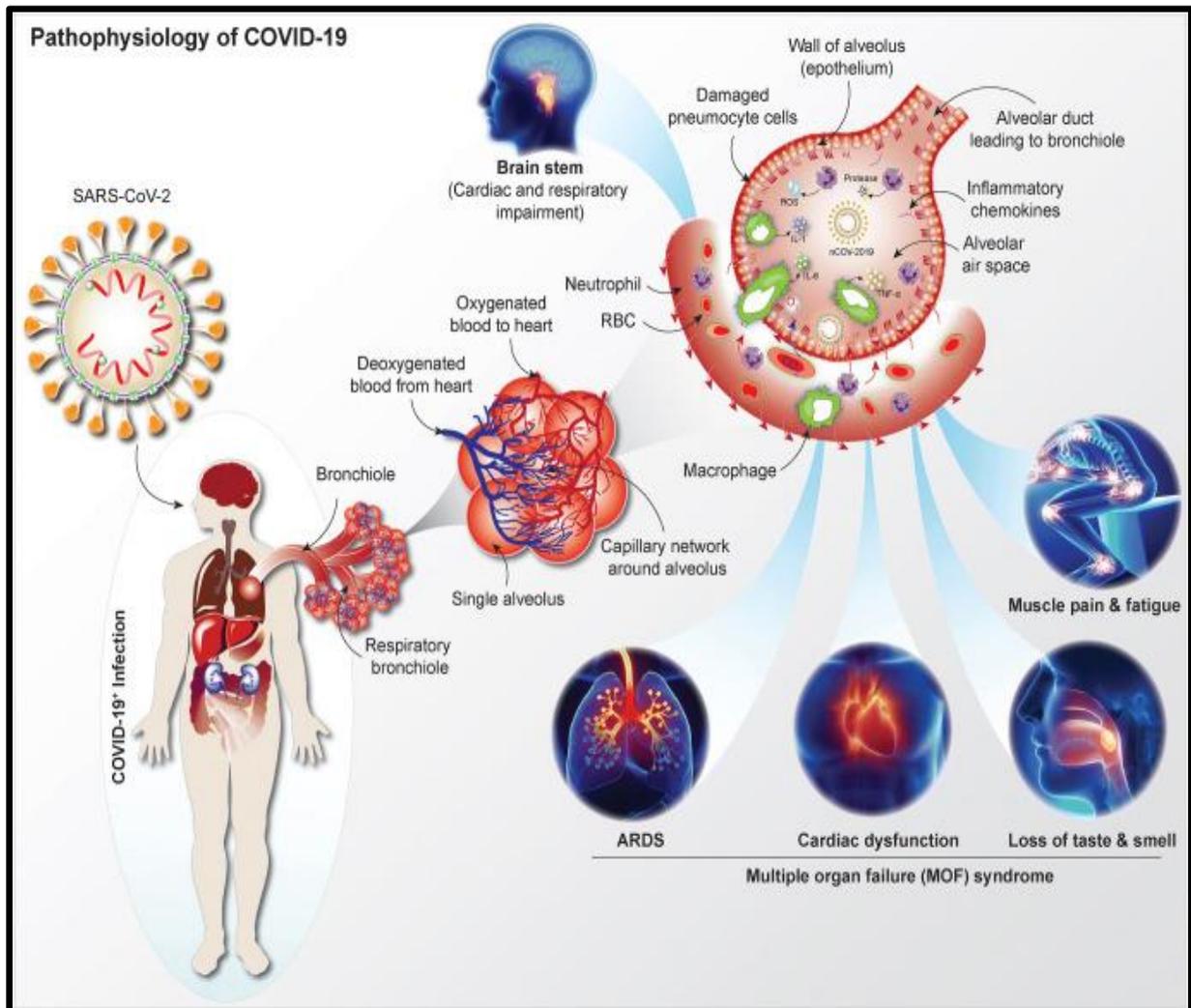


Figure 43 : physiopathologie du SARS-COV-2 (16).

3.9 Diagnostic

Poser rapidement le diagnostic du SARS-COV-2 avec le plus d'exactitude possible est la pierre angulaire du contrôle de la pandémie, néanmoins à l'heure actuelle aucune des méthodes mentionnées ci-dessus n'est parfaite pour diagnostiquer le SARS-COV-2 (82). Au début de la pandémie, le diagnostic du SARS-COV-2 a été basé sur l'historique du patient (anamnèse,

voyage) et sur les analyses biochimiques, mais ces méthodes ne permettaient pas de poser un diagnostic de certitude en raison de l'existence d'autres maladies respiratoires similaires(80). Puis grâce à l'identification des caractéristiques radiographiques spécifiques (infiltrat en verre dépoli) du SARS-COV-2, le TDM (scanner) est devenu un outil indispensable pour orienter le diagnostic, cependant les infrastructures (hôpitaux, laboratoire) qui hébergent ces matériels de pointe se sont très vite retrouvées saturées (matériel onéreux et en nombre limité) (83), c'est pourquoi les scientifiques se sont orientés vers les techniques d'analyse moléculaire qui sont plus précises et permettent d'identifier l'agent viral en cause dans les prélèvements(84) (Figure 44) , les plus utilisées sont ; l'amplification de l'acide nucléique viral « RT PCR » utilisant un ou plusieurs gènes (résultat en 24h), le séquençage de nouvelle génération du génome entier, et les tests antigéniques qui sont les plus rapides et sont vendus comme kit de dépistages (résultat en 30mn) (85).

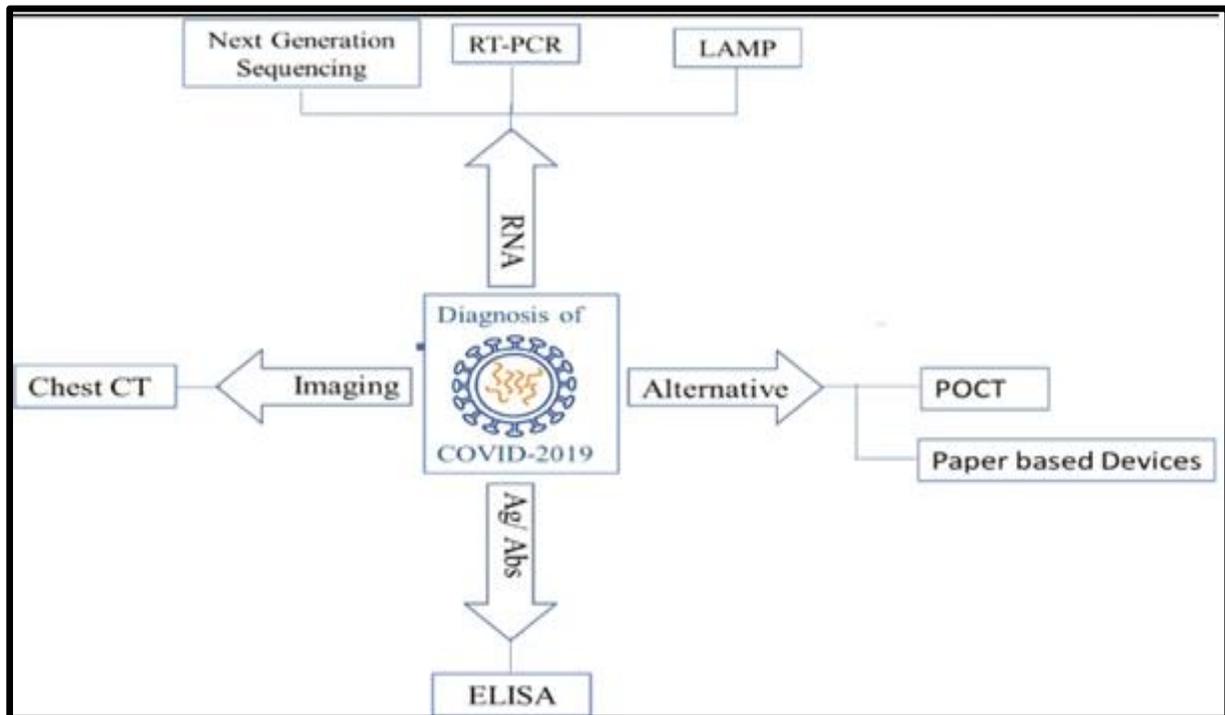


Figure 44 : les différentes méthodes de diagnostic du SARS-COV-2 (84).

Chacune de ces techniques d'analyse possèdent des avantages et des inconvénients qui sont résumés dans le tableau ci-dessous :

Tableau 8 : avantages et inconvénients des méthodes de détection moléculaires du Sars-CoV2 (70).

Méthodes	Principaux avantages	Principaux inconvénients
rRT-PCR real-time reverse transcription polymerase chain reaction	Méthode de référence, sensibilité et spécificité élevées.	TAT long
RT-LAMP Reverse Rranscription Loop- Mediated isothermal amplification	TAT court	Sensibilité plus faible.(moins fiable)
TMA Transcription Mediated Amplification	Automatisation complète, sensibilité élevée, haut débit	TAT long
Test antigénique (ELISA)	TAT court	Moins fiables

TAT : Turn Aroud Time (temps pour obtenir résultat).

3.10 Traitement du SARS-COV-2

À ce jour, il n'existe aucun traitement efficace prouvé contre le SARS-COV-2, mais certains traitements ont produit de bon résultats et ont permis de diminuer le taux de mortalité (86).

Les stratégies thérapeutiques du SARS-COV-2 dépendent essentiellement de la gravité des symptômes ;

3.10.1 Les traitements palliatifs (formes modérées)

Ce sont des traitements symptomatiques qui reposent essentiellement sur la prise en charge de l'hyperthermie en utilisant le PARACETAMOL. Les AINS doivent être évités car pourraient

aggraver la maladie, l'antibiothérapie n'est pas nécessaire pour les cas modérés, elle n'est envisagée que lorsqu'il y a risque de co-infection bactérienne (87,88).

3.10.2 L'élimination virale à l'aide de molécules capables d'interférer avec la réplication virale

Ce sont essentiellement les antiviraux et la chloroquine qui permettent d'inhiber l'entrée du SARS-COV-2 dans la cellule, et ne sont réservés qu'aux formes graves (89).

Le **Remdésivir** et le **Ropinavir** inhibent la réplication virale en interférant avec l'ARN polymérase (86).

La **chloroquine** quant à elle possède un mécanisme d'action multiple, elle permet l'inhibition de la fusion du virus à la surface cellulaire(inhibe la protéine S), le blocage de la réplication virale et a un effet immunomodulateur (87), la chloroquine n'est actuellement plus utilisée en raison de ses effets secondaires graves pouvant entraîner la mort (90).

3.10.3 Les thérapies anti-inflammatoires stéroïdiens (formes graves)

L'utilisation de La dexaméthasone permet vraisemblablement de réduire la durée de la ventilation mécanique et la mortalité causée par le SARS-COV-2 (89) grâce à son activité anti-inflammatoire qui réduit la réponse immunitaire innée et inhibe les cytokines évitant ainsi le choc cytokinique (86).

3.10.4 L'azithromycine (traitement le plus utilisé en Algérie)

L'azithromycine est un antibiotique connu pour ses effets immunomodulateurs et pour sa capacité à induire une réponse antivirale(87), cette dernière semble être lié à l'induction des interférons par augmentation de leurs concentrations dans les cellules bronchiques épithéliales (89).

3.10.5 Les anti-hémostatiques

Le SARS-COV-2 entraîne des complications thrombotiques qui en l'absence d'un traitement adapté entraîne une mortalité plus élevée. Pour éviter ces complications les médecins après avoir confirmé le risque de thrombose (Dosage des D-dimères) prescrivent des médicaments anti hémostatiques, le plus connu est le « **lovenox** » (91).

3.11 Vaccin antiviral

Les vaccins sont une nouvelle arme essentielle dans la lutte contre le SARS-COV-2, Actuellement il existe plus 70 Candidats vaccins en essais clinique sur l'homme, cependant seulement 7 vaccins ont été approuvés par l'OMS (Tableau 9) (92).

Tableau 9 : les 7 vaccins dont l'utilisation est approuvée par l'OMS (92).

Type de vaccin	Nom du vaccin /pays
<p>Vaccin inactivé</p> <p>C'est un virus incapable de se multiplier, il déclenche une réponse immunitaire sans provoquer d'infection</p>	<ul style="list-style-type: none"> • Sinopharm (Chine) • Sinovac (Chine)
<p>Vaccin à vecteur viral</p> <p>Il consiste à utiliser un virus vivant (Adénovirus) rendu inoffensif pour l'homme pour véhiculer une partie précise de l'ADN d'un autre virus (SARS-COV-2) dans les cellules afin de déclencher une réponse immunitaire contre le SARS-CoV-2.</p>	<ul style="list-style-type: none"> • AstraZeneca (Angleterre-Suède) • Johnson&Johnson (USA) • Sputnik V (Russie).
<p>Vaccin à ARN messenger</p> <p>Ils se basent sur l'injection sur l'injection d'un ARN messenger synthétique qui comprend la séquence de la protéine d'intérêt (protéine S pour SARS-COV-2)</p>	<ul style="list-style-type: none"> • Pfizer (USA) • Moderna (USA)

La protéine S est à la base des vaccins actuels contre le SARS-COV-2 dont le but est de générer une réponse immunitaire contre cette protéine. Bien que ces traitements possèdent certains avantages, néanmoins ils ne constituent pas un moyen permanent et efficace pour endiguer la pandémie, preuve en soit, le fait qu'actuellement de nombreux pays font face à une troisième vague du SARS-COV-2, c'est pourquoi la prévention (le port du masque) et la vaccination restent la meilleure solution pour lutter contre le SARS-COV-2 en attendant la découverte d'un traitement ou d'un vaccin à 100% d'efficacité (70,71).

3.12 Impact des variants sur la pathogénie du SARS-COV-2

Les variants du SARS-COV-2 n'entraînent pas pour la plupart des symptômes plus graves, cependant ils ont en commun ; une transmissibilité plus élevée qui fait qu'ils deviennent rapidement dominants au sein des populations, une accumulation d'un nombre élevé de mutations dans la protéine S. Il existe également des signes d'une virulence accrue pour certains variants (variant Alpha et Delta), d'une fréquence de réinfection et d'une résistance accrue à l'action des anticorps monoclonaux et polyclonaux provenant des sérums de convalescence et des vaccins (4).

Le tableau ci-dessous résume les principales caractéristiques virologiques et cliniques de quelques variants du SARS-COV-2 (Tableau 10).

Tableau 10 : Caractéristiques Cliniques et virologiques des principaux variants du SARS-COV-2 (81,82).

Variants	Alpha	Béta	Gamma	Delta	Eta
Pays	Royaume-Uni	Sud d'Afrique	Brésil/Japon	Inde	Nigéria
Déecté en Algérie	Déecté	Non déecté	Non déecté	Déecté	Déecté
Transmission	29% plus transmissible	25% plus transmissible	38 % Plus transmissible	55-97% plus transmissible	plus transmissible
Symptômes	Toux, maux de gorge, fatigue et myalgie plus fréquentes. Perte d'odorat est moins fréquente.	Similaire à ceux du variant Alpha	Similaire à ceux du SARS-COV-2	Toux persistante, maux de gorge, fièvre, écoulement nasal, douleurs musculaire , diarrhée	Similaire à ceux du SARS-COV-2
Sévérité	Augmentation probable du risque de décès : +30-70%	Possible augmentation du risque de décès : +20%	Taux de mortalité plus élevé	Taux de mortalité plus élevé	Pas de donnée disponible

Variants	Alpha	Béta	Gamma	Delta	Eta
Charge virale	Pas de différence significative	Plus élevée	Légèrement Plus élevée	Plus élevée	Pas de différence significative
Détection par test antigénique	Pas de différence significative	Sensibilité du test similaire	Pas de donnée disponible	Pas de différence significative	Pas de différence significative
Impacts potentiels sur l'efficacité des vaccin	Aucun impact significatif Moderna Oxford-AstraZeneca Sinovac	1. Moderna : Réduction de l'activité neutralisante. 2. Novavax et Johnson & Johnson : Efficacité vaccinale plus faible 3. Oxford AstraZeneca : Neutralisation sérologique considérablement réduite par rapport aux souches originales.	Efficacité vaccinale plus faible Environ 80% de protection.	Efficacité vaccinale plus faible Environ 60% de protection pour sinovac Environ 50 % pour Moderna	

A partir de ce tableau nous pouvons constater que le SARS-COV-2 a évolué pendant ces 20 mois de pandémie et continue encore d'évoluer à l'heure actuelle par des mutations qui entraînent l'apparition de nouveaux variants. La détermination des variants et l'analyse des mutations ainsi que l'étude phylogénétique nécessitent une discipline très importante dans le monde scientifique ; la bio-informatique qui est une discipline visant le traitement automatique de l'information biologique et dont l'alignement multiple des séquences constitue une tâche fondamentale pour plusieurs applications en bio-informatique (93).

4 La bio-informatique et la COVID19

4.1 Généralité sur la bio-informatique

La bio-informatique est un domaine interdisciplinaire qui associe l'informatique et les biostatistiques aux sciences biologiques et biomédicales telles que la biochimie, la biologie cellulaire, la biologie du développement, la génétique, la génomique et la physiologie. La bioinformatique est devenue une discipline importante peu après le développement des technologies de séquençage de l'ADN à haut débit dans les années 1970 (94).

L'objectif de la bio-informatique est de faciliter la gestion, l'analyse et l'interprétation des données issues d'expériences biologiques et des études d'observation (93).

4.1.1 Bases de données

Une base de données est un ensemble structure permettant le stockage de grandes quantités d'informations afin de faciliter la récupération et l'utilisation des données. (94)

Pour que les chercheurs puissent bénéficier des données stockées dans une base de données, deux conditions doivent être remplies (94):

- Un accès facile aux informations
- Extraire uniquement les informations nécessaires pour répondre à une question biologique spécifique.

À titre d'exemple, un enregistrement associé à une base de données de séquences nucléiques contient généralement des informations telles que le nom du contact, la séquence d'entrée avec une description du type de molécule, le nom scientifique de l'organisme source à partir duquel elle a été isolée et, souvent, des citations bibliographiques associées à la séquence (93).

4.1.2 Banques de Données Biologiques

Les premières banques de données biologiques sont apparues au début des années 80 sous l'initiative de quelques chercheurs dont la principale mission est de rendre publiques les séquences qui ont été déterminées. Les données biologiques stockées dans ces banques sont des séquences primaires d'ADN, d'ARN et de protéines. Les données peuvent être soumises et consultées par l'intermédiaire du Web. Les séquences stockées dans ces banques sont obtenues

de plusieurs manières différentes : isolées à partir d'une cellule, déduites à partir de la séquence nucléique par simple traduction (ARN ou protéines) ou encore par génie génétique (94).

4.1.2.1 Les Banques de Séquences Nucléiques

4.1.2.1.1 EMBL

C'est une banque européenne créée en 1980 et financée par l'EMBO (European Molecular Biology Organization), elle est aujourd'hui diffusée par l'EBI (European Bioinformatics Institute, Cambridge, UK). L'EMBL-EBI offre un accès gratuit à des applications populaires d'analyse de séquences bio-informatiques ainsi qu'à un moteur de recherche textuel complet doté de puissantes capacités de référencement croisé et d'extraction de données (95).

4.1.2.1.2 GenBank

Créée en 1982 par la société IntelliGenetics et diffusée maintenant par le NCBI (National Center for Biotechnology Information, Los Alamos, US), Elle est soutenue par le NIH (National Institute of Health et possède plus de 50 millions séquences stockées (96).

4.1.2.1.3 DDBJ (Dna Data Bank)

Elle est créée en 1986 et diffusée par le NIG (National Institute of Genetics, Japon). La Banque de données ADN du Japon est une base de données publique de séquences nucléotidiques Depuis 1987, la DDBJ recueille des séquences nucléotidiques annotées en collaboration avec la GenBank et l'EMBL (97).

4.1.2.2 Les Banques de Séquences Protéiques

4.1.2.2.1 PIR-NBRF

Elle est créée en 1984 par la NBRF (National Biomedical Research Foundation). Elle est maintenant un ensemble de données issues du MIPS (Martinsried Institute for Protein Sequences, Munich, Allemagne) et de la banque japonaise JIPID (Japan International Protein Information Database) (98).

4.1.2.2.2 SWISSPROT

Elle est créée en 1986 à l'Université de Genève et maintenue depuis 1987 dans le cadre d'une collaboration, entre cette université (via ExPASy, Expert Protein Analysis System) et l'EBI. Elle regroupe aussi des séquences annotées de la banque PIRNBRF ainsi que des séquences codantes traduites de l'EMBL (99).

4.1.3 LES GENOME BROWSERS

Ils correspondent à différentes bases de données qui permettent d'accéder aux données du génome humain (et de celui d'autres espèces) à l'aide d'une interface graphique.(100). Les plus populaires sont :

- ❖ **Ensembl**: est un projet commun entre l'EBI (European Bioinformatics Institute) et le centre Sanger (100).
- ❖ **NCBI** (National Cancer for Biology Information) : est un site américain hébergeant une banque de gènes, d'ARNm et de protéines qui est régulièrement complétée par des laboratoires du monde entier et par des centres de séquençage (91).
- ❖ **UCSC** (University of California Santa Cruz) : Il s'agit d'un site web qui offre un accès aux données relatives aux séquences génomiques d'une variété d'espèces (100).

Les séquences disponibles sur les Génome Browser sont au format **FASTA**, c'est un fichier texte utiliser pour stocker des séquences biologiques de nature nucléique ou protéique (101).

4.2 Rôle de la bio-informatique dans l'étude de la physiopathologie du SARS COV2 et variants

L'exploitation des outils et techniques de la bio-informatique a permis d'interpréter avec succès l'architecture génomique virale du SARS-COV-2 dans un délai très court.

La bio-informatique a non seulement permis de séquencer le génome du SARS-COV-2 mais aussi d'analyser correctement les erreurs de séquençage, les relations évolutives et les variations génétiques par l'application d'un algorithme d'alignement des séquences nucléiques et protéiques. Ces données sont essentielles pour le développement d'un vaccin contre le SARS-COV-2 (102).

4.3 Alignement des séquences

L'alignement de séquences est une façon de représenter deux ou plusieurs séquences (ADN, ARN ou protéines) les unes sous les autres afin d'identifier les zones de concordance. Ces alignements sont réalisés par des programmes informatiques. Il existe deux classes d'alignement (101).

4.3.1 Alignement de deux séquences

Un alignement de deux séquences (appelé souvent ‘Alignement deux à deux’) est une mise en correspondance entre les résidus avec une possible insertion des espaces « gaps » afin d’obtenir des séquences de longueur égales (Figure 45) (103).

Trois situations sont possibles pour une position donnée de l’alignement (104):

- ❖ Les caractères sont les mêmes : identité
- ❖ Les caractères ne sont pas les mêmes : Substitution
- ❖ L’une des positions est un gap (espace) : Insertion/Délétion
- ❖ Il existe plusieurs algorithmes d’alignement, les plus connus sont :
 - Clustal omega, W, X
 - Muscle « *Multiple Sequence Comparison by Log-Expectation* »(104).

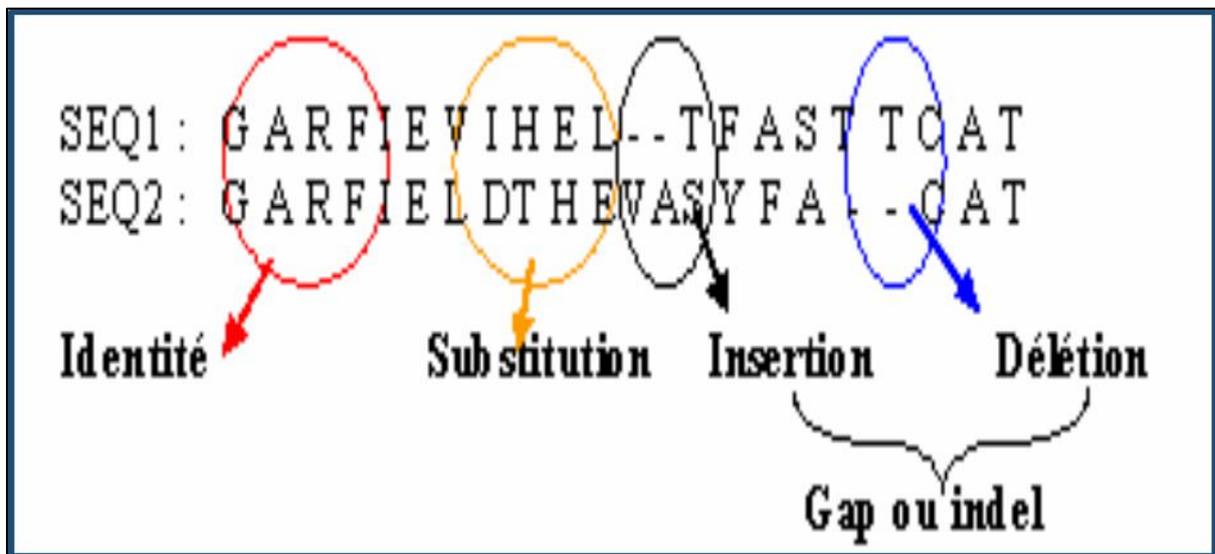


Figure 45 : alignement de deux séquences protéiques (101).

4.3.2 L’Alignement Multiple des Séquences

L’alignement multiple permet d’aligner globalement des séquences et de détecter les régions qui ont été conservées au travers de l’évolution (Figure 46) (104).

Les différents objectifs visés par un alignement multiple de séquences sont (101):

- Trouver des caractéristiques communes à une famille de protéines.
- Caractériser les régions conservées et les régions variables.
- Construire l’arbre phylogénétique de séquences homologues(101).

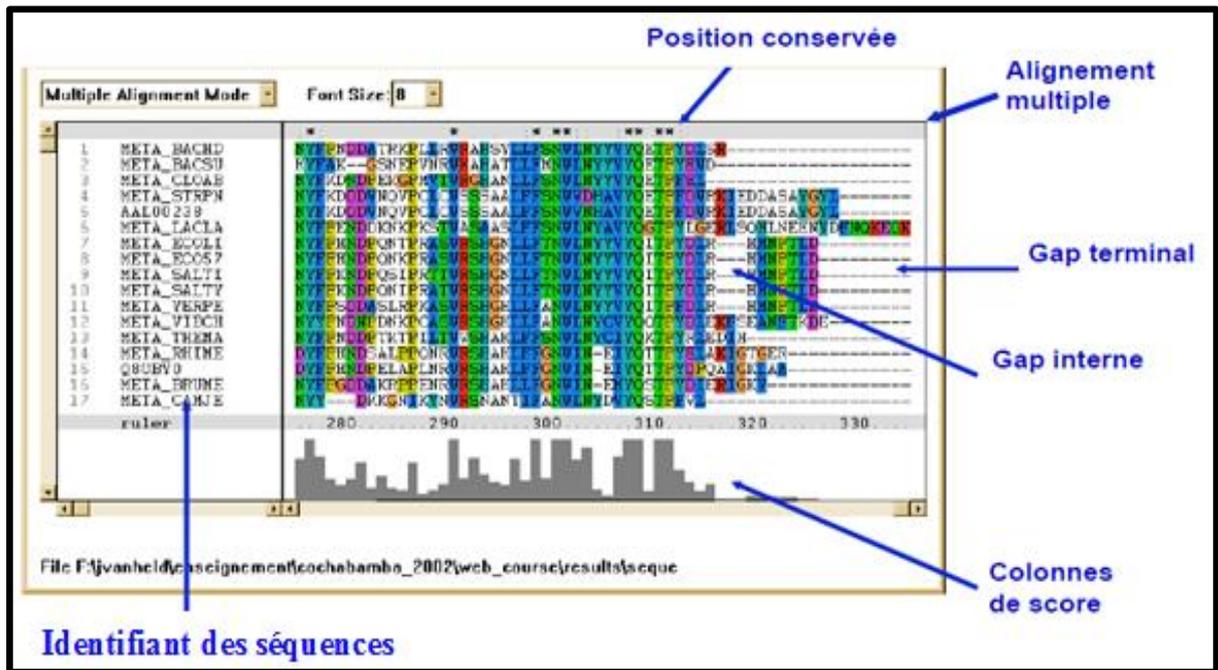


Figure 46 : exemple d'un alignement multiple de séquences protéiques (104).

Partie
pratique

5 Matériels et méthodes

Dans ce travail de recherche nous avons décidés d'étudier et de faire l'alignement des variants détectés en Algérie ainsi que de quelques variants les plus contagieux dans le monde à ce jour :

- Variants détectés en Algérie ; « Alpha, Delta, Eta »
- Variants les plus contagieux dans le monde ; « Béta, Gamma ».

Le but de l'alignement des séquences est de mettre en valeur l'importance des mutations et leurs impacts sur la virulence du virus ainsi que sur la symptomatologie et l'efficacité des vaccins et traitements actuels (101).

5.1 Acquisition des séquences nucléotidiques et protéiques

Dans ce travail nous avons décidés d'utiliser un seul moteur de recherche génomique qui est le NCBI, cette décision est principalement fondée sur une appréciation personnelle qui est la facilité de recherche des articles et séquences voulues.

L'acquisition des séquences passe par plusieurs étapes qui sont présentées ci-dessous :

5.1.1.1 Rechercher sur NCBI la séquence d'intérêt

A l'aide du moteur de recherche intégré, on recherche la séquence nucléotidique souhaitée. Dans notre cas on recherche la séquence protéique de la glycoprotéine S du SARS-COV-2. On sélectionne la catégorie « Nucléotide » dans le bandeau de recherche, puis écrire le nom du gène ou de la protéine (ici « Spike protein Wuhan ») (Figure 1).

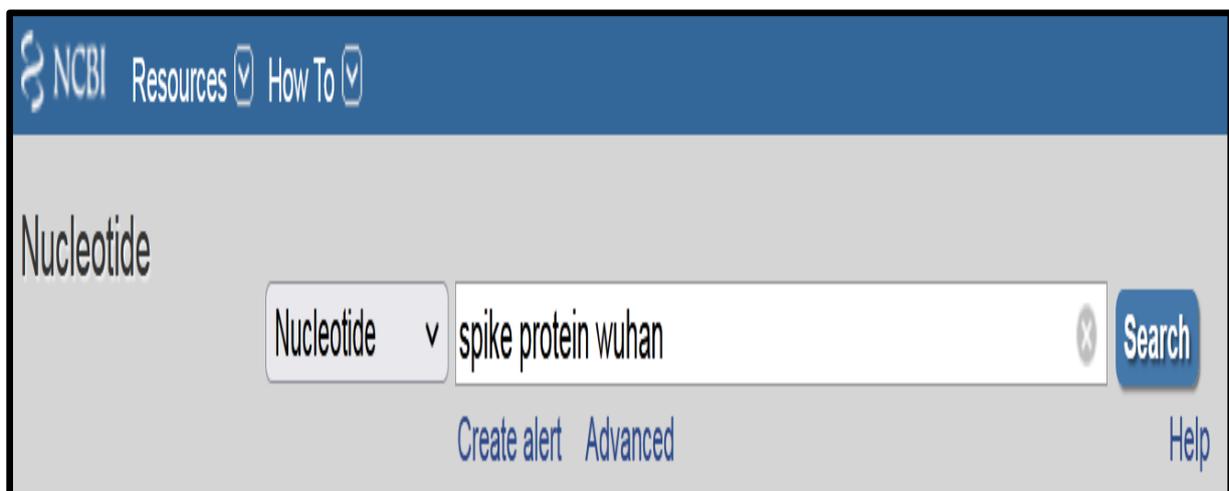


Figure 1 : Moteur de recherche du site NCBI.

5.1.1.2 Identifier la séquence souhaitée

Trier les résultats par espèces et par le virus recherché (ici « severe acute respiratory syndrom ») puis sélectionner la séquence souhaitée (Figure 2).

The screenshot shows the NCBI Nucleotide search interface. The search term is "spike protein wuhan". The results are filtered by species, with "Viruses (391)" selected. The top organisms list includes "Severe acute respiratory syndrome coronavirus 2 (6)".

Species List:

- Animals (21)
- Plants (7)
- Bacteria (4)
- Viruses (391)**
- Customize ...

Top Organisms [Tree]:

- Porcine epidemic diarrhea virus (190)
- Human rotavirus A (47)
- Swine acute diarrhea syndrome coronavirus (36)
- Porcine deltacoronavirus (16)
- Bat SARS-like coronavirus (11)
- Homo sapiens (10)
- Severe acute respiratory syndrome coronavirus 2 (6)**
- mammalian orthoreovirus (6)
- Mus musculus (5)
- Feline coronavirus (4)
- Swine acute diarrhea syndrome related coronavirus (4)
- Malus domestica (4)
- Coronavirus AcCoV-JC34 (2)
- Bat SARS-like coronavirus WIV1 (2)

Figure 2 : filtrage des résultats de recherche

5.1.1.3 Choisir le résultat de recherche qui correspond à notre recherche

Une liste de plusieurs résultats apparait, on sélectionne le résultat qui correspond à notre recherche. Dans notre cas le choix de la séquence s'est fait par rapport à la longueur de la séquence du génome (Figure 3).

The screenshot shows the NCBI Nucleotide search interface. The search query is "ND \"Severe acute respiratory syndrome coronavirus 2\"[porgn: __txid2697049]". The results are filtered to show 6 items, with the first three items listed. The first item, "Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome", is highlighted with a red box and a large red number "4" next to it. This item has a length of 29,903 bp linear RNA and accession number NC_045512.2. The second item is "Severe acute respiratory syndrome coronavirus 2 isolate WIV07, complete genome" with a length of 29,857 bp linear RNA and accession number MN996531.1. The third item is "Severe acute respiratory syndrome coronavirus 2 isolate WIV06, complete genome" with a length of 29,854 bp linear RNA and accession number MN996531.1. The interface includes filters for Viruses, Molecule types, Source databases, and Sequence Type.

Figure 3 : sélection du résultat de recherche voulu.

Une fois le lien ouvert, on voit apparaître la séquence globale du génome du SARS-COV- 2 qui est divisé en plusieurs séquences qui correspondent aux domaines qui constituent le génome viral (ORF1a, ORF1b, protéine S, ORF3a, protéine E ...etc.) (Figure 4).

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

NCBI Reference Sequence: NC_045512.2

[FASTA](#) [Graphics](#)

Identifiant séquence NCBI	Type + longueur du génome	Date publication
LOCUS NC_045512	29903 bp ss-RNA linear	VRL 18-JUL-2020
DEFINITION	Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1; complete genome.	
ACCESSION	NC_045512	
VERSION	NC_045512.2	
DBLINK	BioProject: PRJNA485481	
KEYWORDS	RefSeq.	
SOURCE	Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)	
ORGANISM	Severe acute respiratory syndrome coronavirus 2 Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes; Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.	

↑
Souche

Figure 4 : page du génome complet du SARS-COV-2 sur le site NCBI.

5.1.1.4 Faire apparaître la séquence d'intérêt

Faire défiler la page jusqu'à trouver la section concernant la protéine S puis sélectionner l'identifiant de la séquence (ici « YP_009724390.1 ») (Figure 5).

gene	21563..25384	Localisation du gène S
	/gene="S"	
	/locus_tag="GU280_gp02"	
	/gene_synonym="spike glycoprotein"	
	/db_xref="GeneID:43740568"	
CDS	21563..25384	
	/gene="S"	
	/locus_tag="GU280_gp02"	
	/gene_synonym="spike glycoprotein"	
	/note="structural protein; spike protein"	
	/codon_start=1	
	/product="surface glycoprotein"	
	/protein_id="YP_009724390.1"	Identifiant NCBI protéine S
	/db_xref="GeneID:43740568"	
	/translation="MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYDPKVFR SSVLHSTQDLFLPFFSNVTWFWFAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIIR GWIFGTTLDKSTOSLLIVNNAATNVVIVKVFQFCNDPFLGVYYHKNNKSWMESEFRVY	

5

Figure 5 : identification NCBI de la séquence protéique de la protéine S.

5.1.1.5 Utiliser le lien FASTA pour faire apparaître la séquence protéique

surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

NCBI Reference Sequence: YP_009724390.1

[Identical Proteins](#) [FASTA](#) [Graphics](#)

[Go to:](#)

6

Longueur séquence

LOCUS YP_009724390 1273 aa linear VRL 18-JUL-2020

DEFINITION surface glycoprotein [Severe acute respiratory syndrome coronavirus 2].

ACCESSION YP_009724390

VERSION YP_009724390.1

DBLINK BioProject: [PRJNA485481](#)

DBSOURCE REFSEQ: accession [NC_045512.2](#)

KEYWORDS RefSeq.

SOURCE Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

ORGANISM [Severe acute respiratory syndrome coronavirus 2](#)

Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes; Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.

Figure 6 : choix du format FASTA

5.1.1.6 Apparition de la séquence de la protéine S

La séquence affichée est une succession de lettres latines dont chaque lettre correspond à un acide aminé pour les séquences protéiques, et à des bases nucléiques pour les séquences nucléiques (Figure 7)

surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

NCBI Reference Sequence: YP_009724390.1

[GenPept](#) [Identical Proteins](#) [Graphics](#)

>YP_009724390.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

```
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNATNVVIVKCEFCNDPF
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLT PGDSSSGWTAGAAAYVGYLQPRTFLLKYN
ENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV
YAWNKRKISNCVADYSVLYNSASFSTFKCYGVSPTKLNLDLCTNRYADSFVIRGDEVQRQIAPGQTGKIAD
YNYKLPDDFTGCVIAWNSNNLDSKVGNNYNYLYRFRKSNLKPFFERDISTEIQAGSTPCNGVEGFNCYF
PLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNGLTGTGVLTESNKKFL
PFQQFGRDIADTTDAVRDPQTLLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT
PTWRVYSTGSNVFQTRAGCLIGAEHVNSYECDIPIGAGICASYQTQTNSPRRARSVASQSI IAYTMSLG
AENSVAYSNNIAIPTNFTISVTEILPVSMTKTSVDCTMYICGDSTECSNLLQYGSFCTQLNRALTGI
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSPKSKRSFIEDLLENKVTLADAGFIKQYGDC
LGDIAARDLICAQKFNGLTVLPPLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG
VTQNVLYENQKLIANQFNSAIGKIQDLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDI
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM
SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAI CHDGAHF PREGVFSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVA
KNLNE SLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKCCSCGSCCKFDEDD
SEPV LKGVKLHYT
```

Figure 7 : apparition de la séquence au format FASTA

5.1.1.7 Obtention des séquences de la protéine S des variants

Les étapes précédemment citées sont réitérées afin d'obtenir les séquences des variants G614, Alpha, Béta, Eta, Delta, Gamma.

Le tableau ci-dessous contient les numéros d'accessions (ID séquence) des séquences de la protéine S utilisées dans ce mémoire.

Tableau 1 : numéro d'accession des séquences de la protéine S.

Souches	ID séquence	
	Protéine S	Génome Complet
Ancestrale	YP_009724390.1	NC_045512.2
G614	QTA38988.1	QNH88943.1
Alpha	QSV99683.1	MW735429.1
Béta	QWA53387.1	MZ314998.1
Delta	QWP92470.1	MZ377116.1
Eta	QWM26437.1	MZ362451.1
Gamma	QXU68227.1	MZ611957.1

5.2 L'Alignement Multiple des Séquences protéiques

Il existe une multitude de logiciels et sites en ligne qui permettent de réaliser un alignement multiple des séquences protéiques ou nucléique, le plus connu est le Clustal OMEGA (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) . Cependant l'outil Clustal ne permet pas l'édition et la visualisation des résultats, ainsi afin de faciliter la compréhension des résultats obtenus et les faire correspondre à l'objectif principale de ce travail qui est d'étudier les mutations observées dans les variants du SARS-COV-2 ainsi que leurs impacts sur la virulence du virus, nous avons décidé d'utiliser le logiciel d'alignement « JALVIEW » qui est doté d'une interface très simple et facilement configurable en fonction du résultat d'alignement voulu.

Le logiciel JALVIEW est disponible gratuitement sur le site officiel du développeur « <https://www.jalview.org/> », nous allons expliquer les différentes étapes de l'utilisation de ce logiciel pour réaliser l'alignement multiple, pour cela suivre les étapes décrites ci-dessous :

- Importer les séquences ; appuyer sur « File » puis « Input Alignment » puis « From Textbox » (Figure 8).

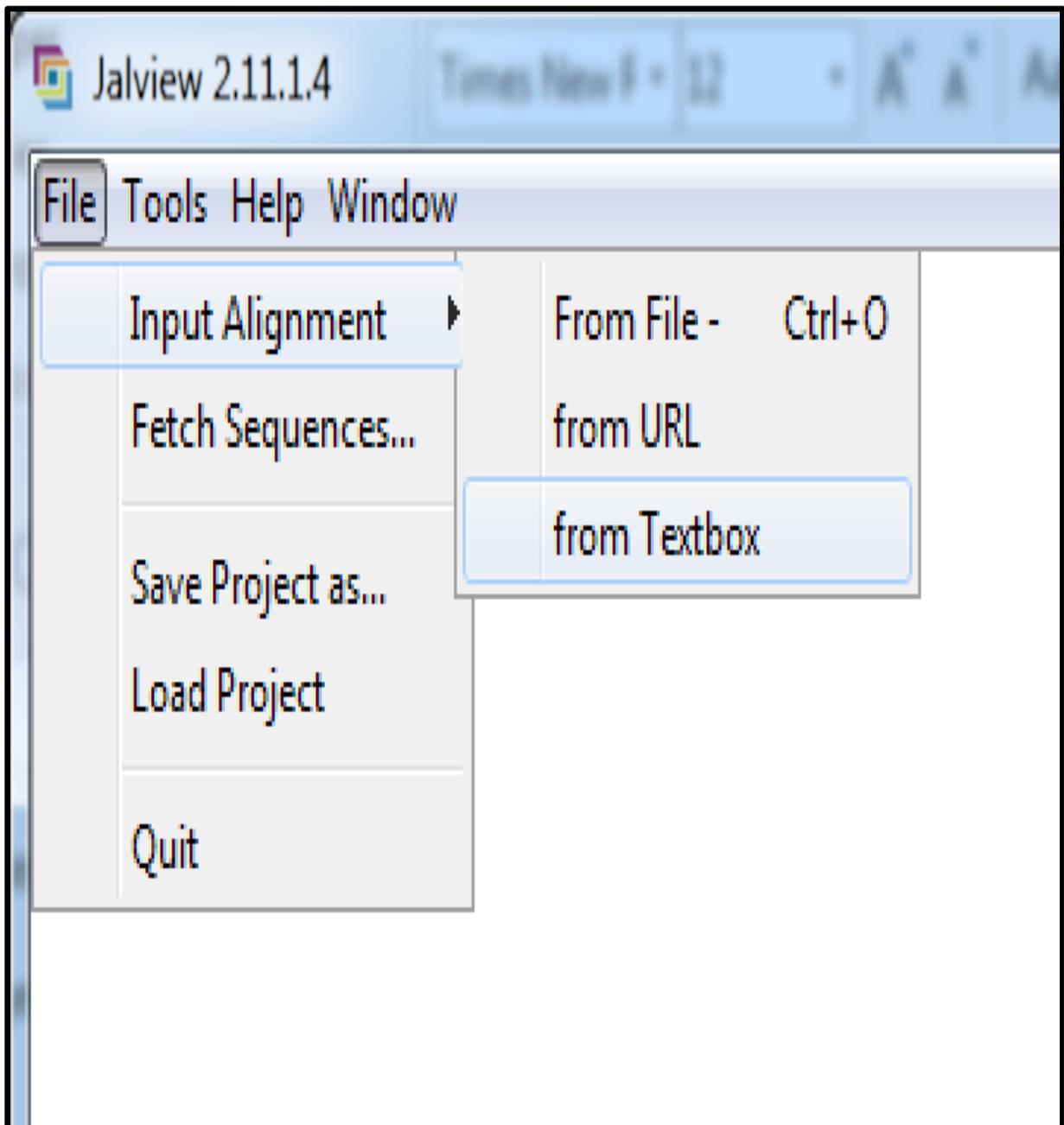


Figure 8 : importer les séquences dans le logiciel JALVIEW.

- Une nouvelle console s'ouvre dans laquelle on copie les séquences au format FASTA (Figure 9).

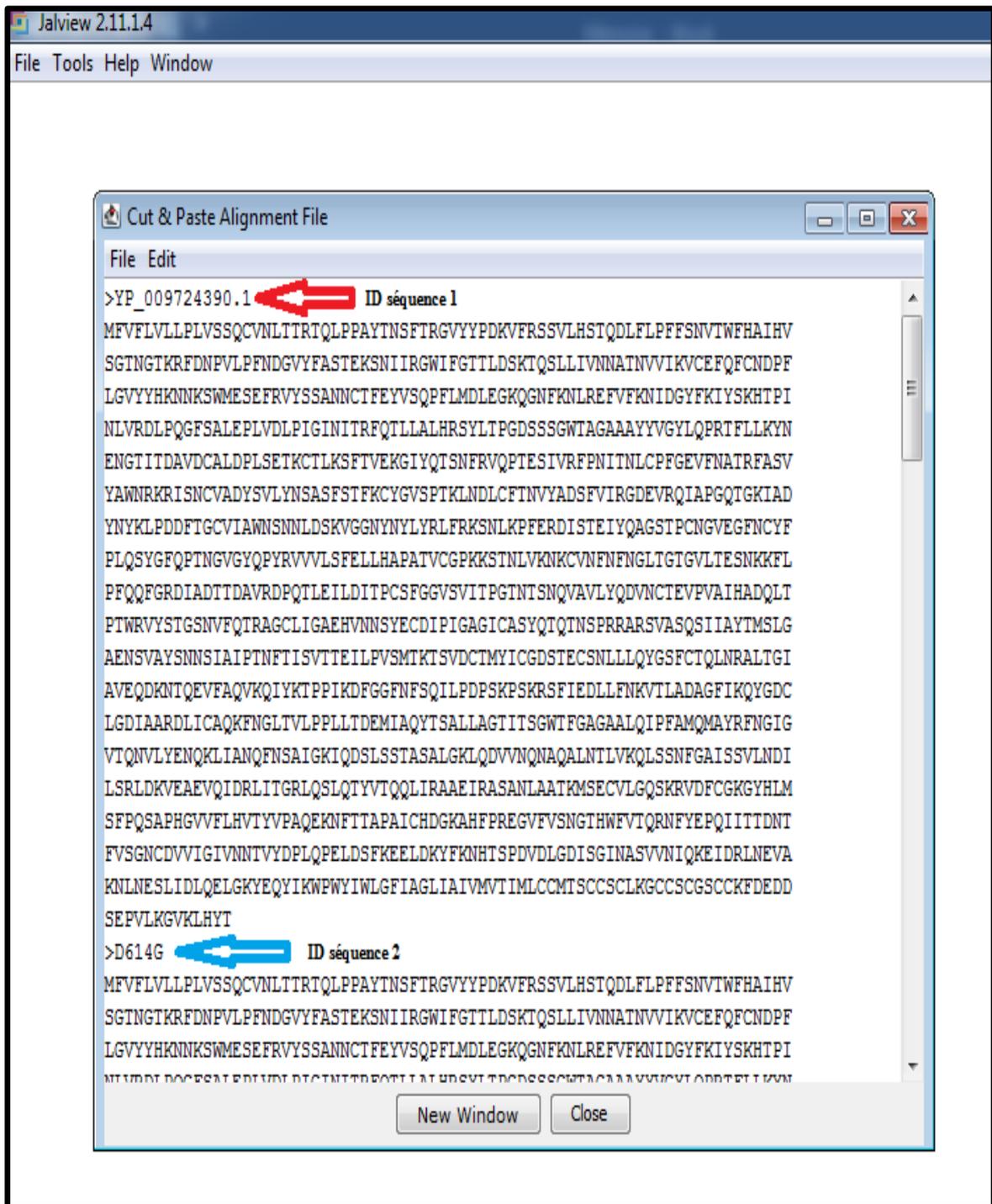


Figure 9 : Séquences copiées dans la console du logiciel.

- Appuyer sur « **New window** » et une nouvelle fenêtre s'affiche qui contient les séquences non alignées disposées l'une en dessous de l'autre (Figure 10).

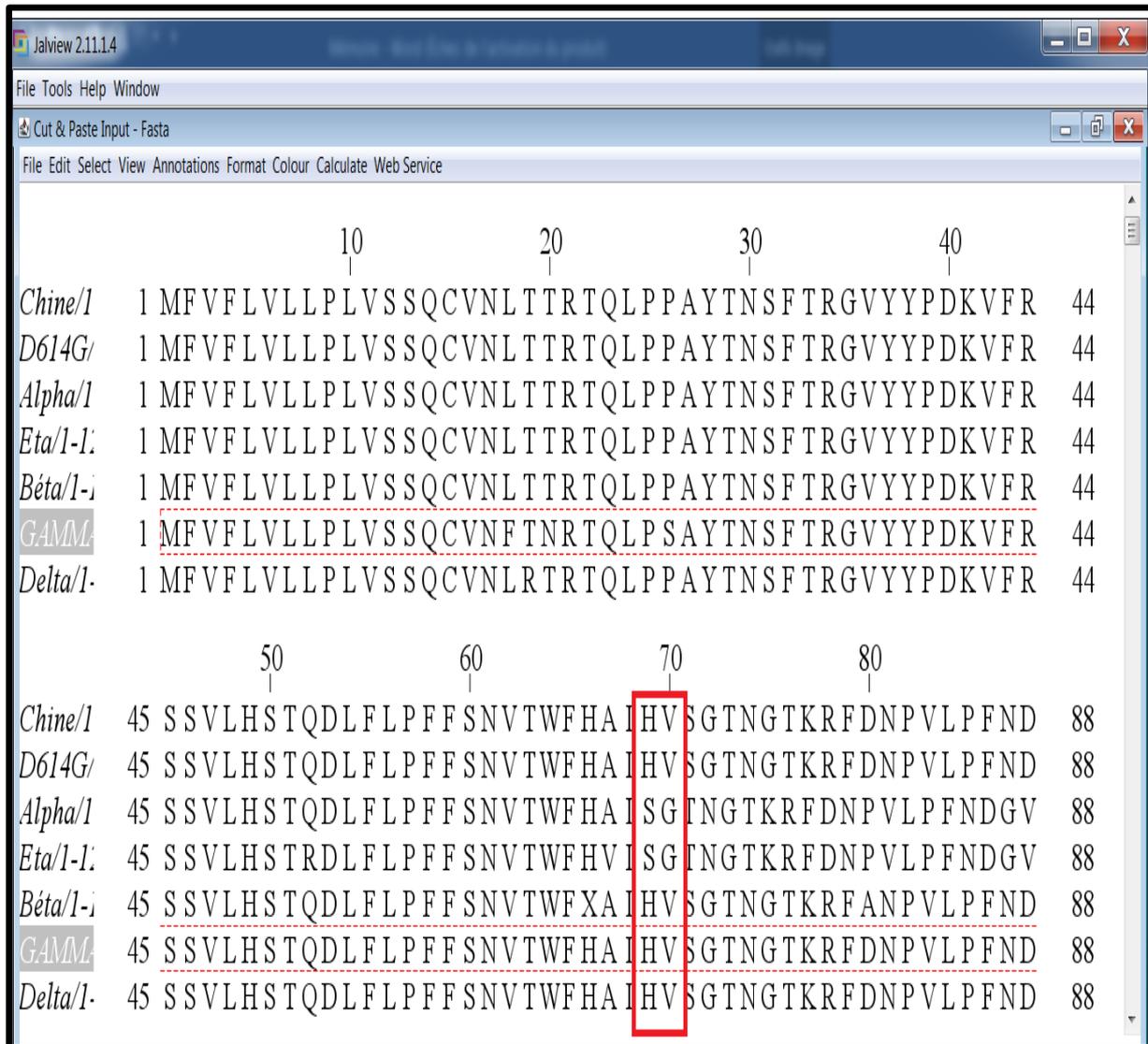


Figure 10 : séquences non encore alignées par l'algorithme d'alignement.

- Pour exécuter l'algorithme d'alignement multiple, allez à l'onglet « **Web Service** » puis « **Alignment** » puis sélectionner le type le langage d'alignement, dans notre cas nous avons utilisé le « **Muscle with default** » car il a à la fois une meilleure précision et une meilleure vitesse que le CLUSTALW OU T-COFFEE pour l'alignement des courtes séquences.

6 Résultats et discussions

6.1 Résultats de l'alignement des séquences

La figure ci-dessous montre le résultat de l'alignement multiple après le traitement des séquences par l'algorithme d'alignement (Figure 11).

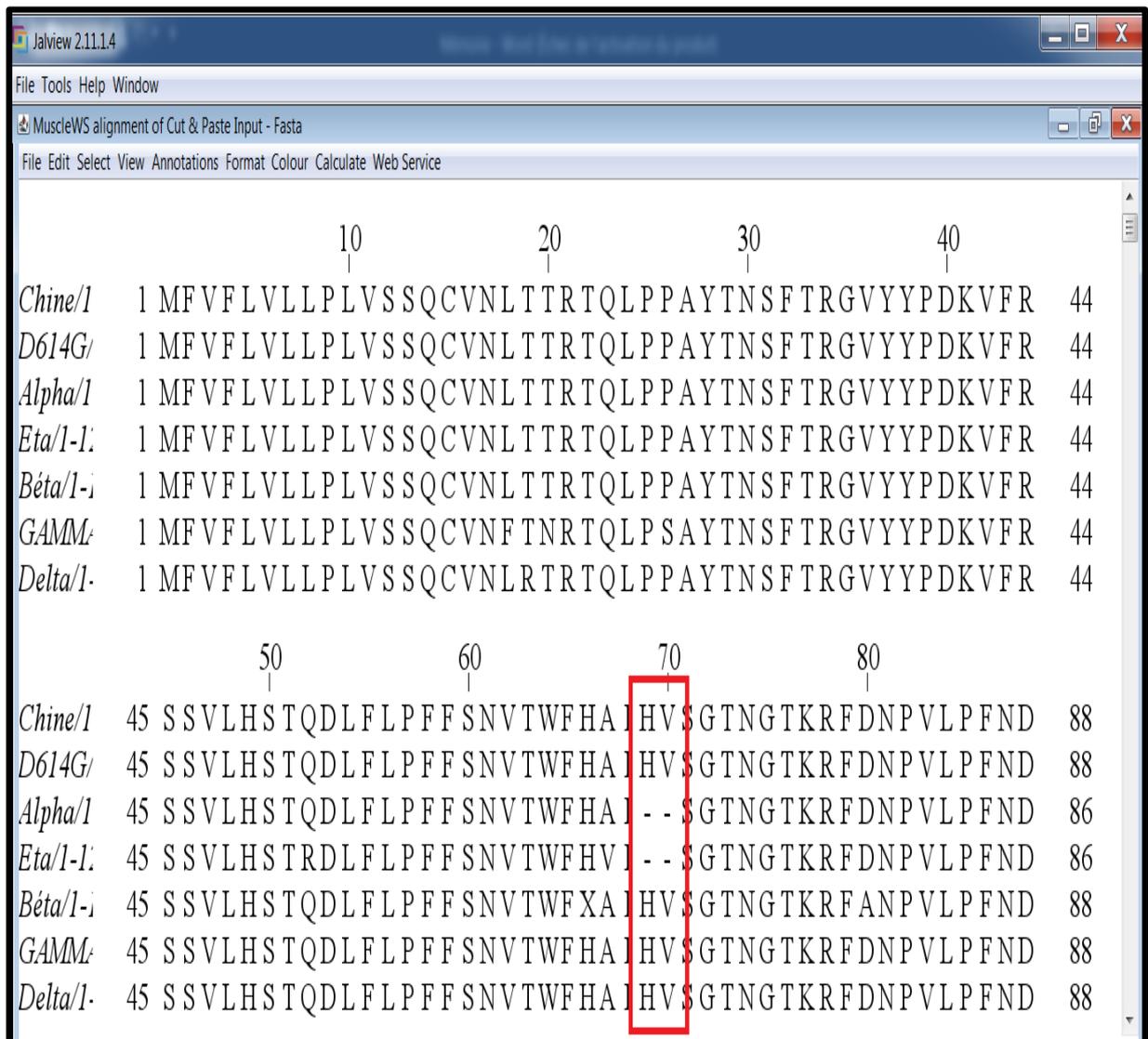


Figure 11 : résultat de l'alignement multiple des séquences protéiques.

La figure 11 montre le résultat du traitement des séquences par l'algorithme d'alignement multiple et ainsi nous pouvons constater la différence entre la figure 10 et 11 qui consiste en l'introduction des « Gap » afin de maximiser le nombre de coïncidence entre les acides aminés dans les séquences.

Pour afficher les différences entre les séquences, on doit sélectionner « Color » puis choisir la coloration souhaitée ; à titre d'exemple la coloration Clustal colore les acides aminés en fonction de leur propriétés chimique (hydrophobes, hydrophiles). Dans notre travail nous avons choisis la coloration « *pourcentage identity* » (pourcentage d'identité) car nous essayons principalement de mettre en évidence les différences du génome entre la souche initiale du SARS-COV-2 et ses variants plus particulièrement les mutations survenues au niveau de la protéine S dans les différents variants du SARS-COV-2 (Figure 12).



Figure 12 : résultat de l'alignement multiple des séquences protéiques.

La figure 12 représente le résultat de l'alignement multiple après avoir appliqué la coloration « *pourcentage d'identité* » pour afficher les régions homologues et les régions ayant subi des mutations, ainsi le résultat de la figure est interprété comme suit :

- ✚ La couleur bleue représente la similitude entre les séquences,
- ✚ Le blanc indique les différences entre les séquences
- ✚ Le symbole « - » et la lettre X indiquent une délétion.

La longueur des séquences de la protéine S des variants est plus courtes que celle de la souche initiale à l'exception du variant Gamma qui ne comprend aucune délétion (Tableau 1) (Figure 13).

Tableau 2 : comparaison de la longueur des séquences protéiques de la protéine S.

Souches	Longueur séquence
Souche initiale	1273 AA (acides aminés)
Variant G614	1273 AA
Variant Alpha	1270 AA
Variant Béta	1272 AA
Variant Gamma	1273 AA
Variant Eta	1270 AA
Variant Delta	1271 AA

		1270	
<i>Chine/</i>	1261	SEPVLKGVKLHYT	1273
<i>D614G/</i>	1261	SEPVLKGVKLHYT	1273
<i>Alpha/</i>	1258	SEPVLKGVKLHYT	1270
<i>Eta/</i>	1260	SEPVLKGVKLHYT	1272
<i>Béta/</i>	1259	SEPVLKGVKLHYT	1271
<i>GAMMA/</i>	1258	SEPVLKGVKLHYT	1270
<i>Delta/</i>	1261	SEPVLKGVKLHYT	1273

Figure 13 : longueur des séquences protéiques dans le logiciel **JALVIEW**.

A l'aide du logiciel de l'alignement multiple, nous pouvons visualiser les différences entre les séquences protéiques des variants par rapport à la souche initiales ainsi que la position des mutations et le type de mutations, nous allons étudier les mutations les plus dangereuses particulièrement celles qui ont un impact sur la virulence du virus (transmissibilité, évasion immunitaire, résistance aux vaccins et médicaments)

6.1.1 Délétion H69/V70

		70	80	90	
<i>Chine/</i>	65	FHA	HVSGTNGTKRFDNPVLPFNDGVYFASTE		96
<i>D614G/</i>	65	FHA	HVSGTNGTKRFDNPVLPFNDGVYFASTE		96
<i>Alpha/</i>	65	FHA	--SGTNGTKRFDNPVLPFNDGVYFASTE		94
<i>Eta/</i>	65	FHV	--SGTNGTKRFDNPVLPFNDGVYFASTE		94
<i>Béta/</i>	65	FXA	HVSGTNGTKRFANPVLPFNDGVYFASTE		96
<i>GAMMA/</i>	65	FHA	HVSGTNGTKRFDNPVLPFNDGVYFASTE		96
<i>Delta/</i>	65	FHA	HVSGTNGTKRFDNPVLPFNDGVYFASTE		96

Délétion H69/V70

Figure 14 : Délétion H69/V70 détectée dans les variants Alpha et Eta.

La figure 14 indique la présence de la délétion **H69/V70** dans deux variants ; Alpha et Eta. À l'heure actuelle sur les 21 variants connus, seulement un seul autre variant « B.1.620 » possède cette mutation.

La délétion **H69/V70** entraîne une perte de détection du Gène S (SGTF) ainsi qu'une augmentation de la transmissibilité et une diminution de la neutralisation du virus par les anticorps (105).

D'après une étude réalisée par des chercheurs japonais (*Takano et al.*), la délétion **V69/H70** à elle seule entraîne un prolongement de l'inflammation (dure plus longtemps) , une augmentation de la charge virale ainsi qu'une aggravation de la maladie (106).

6.1.2 Mutation K417N/T

		410		420		430																										
<i>Chine/</i>	409	Q	I	A	P	G	Q	T	G	K	A	D	Y	N	Y	K	L	P	D	D	F	T	G	C	V	I	A	W	N	S	N	439
<i>D614G/</i>	409	Q	I	A	P	G	Q	T	G	K	A	D	Y	N	Y	K	L	P	D	D	F	T	G	C	V	I	A	W	N	S	N	439
<i>Alpha/</i>	406	Q	I	A	P	G	Q	T	G	K	A	D	Y	N	Y	K	L	P	D	D	F	T	G	C	V	I	A	W	N	S	N	436
<i>Eta/</i>	406	Q	I	A	P	G	Q	T	G	K	A	D	Y	N	Y	K	L	P	D	D	F	T	G	C	V	I	A	W	N	S	N	436
<i>Béta/</i>	408	Q	I	A	P	G	Q	T	G	N	A	D	Y	N	Y	K	L	P	D	D	F	T	G	C	V	I	A	W	N	S	N	438
<i>GAMMA/</i>	409	Q	I	A	P	G	Q	T	G	T	A	D	Y	N	Y	K	L	P	D	D	F	T	G	C	V	I	A	W	N	S	N	439
<i>Delta/</i>	407	Q	I	A	P	G	Q	T	G	K	A	D	Y	N	Y	K	L	P	D	D	F	T	G	C	V	I	A	W	N	S	N	437

K417N/T

Figure 15 : Mutation K417N présente dans le variant Béta et K417T dans le variant Gamma.

La figure 15 montre la présence de la mutation **K417N** dans le variant Béta et la mutation **K417T** dans le variant Gamma, actuellement seuls ces deux variants possèdent la mutation.

La mutation **K417N/T** est impliquée dans des changements structuraux de la protéine **S** et l'évasion immunitaire. Elle semble aussi diminuer l'affinité de la protéine **S** pour les récepteurs **ACE2** qui d'après des récentes études semble être due à la rupture de la liaison hydrogène entre la lysine 417 et l'acide aspartique 101 provoquée par la mutation **N417**, cependant cette diminution est compensée par la présence de la mutation **N510Y** (107,108).

6.1.3 Mutation L452R

	440	450	460						
<i>Chine/</i>	440	NLDSKVG	GNYNYLYR	LF	FRKSNL	KPFERD	IST	470	
<i>D614G/</i>	440	NLDSKVG	GNYNYLYR	LF	FRKSNL	KPFERD	IST	470	
<i>Alpha/</i>	437	NLDSKVG	GNYNYLYR	LF	FRKSNL	KPFERD	IST	467	
<i>Eta/</i>	437	NLDSKVG	GNYNYLYR	LF	FRKSNL	KPFERD	IST	467	
<i>Béta/</i>	439	NLDSKVG	GNYNYLYR	LF	FRKSNL	KPFERD	IST	469	
<i>GAMMA/</i>	440	NLDSKVG	GNYNYLYR	LF	FRKSNL	KPFERD	IST	470	
<i>Delta/</i>	438	NLDSKVG	GNYNYR	YR	LF	FRKSNL	KPFERD	IST	468

L452R

Figure 16 : Mutation L452R détectée dans le variant Delta.

La figure 16 indique la détection de la mutation **L452R** dans le variant Delta, récemment la mutation a aussi été détecté dans les variants Kappa, Epsilon et Lambda.

La mutation **L452R** semble être impliquer dans l'augmentation de la transmissibilité et l'évasion immunitaire(109).

Selon « *Saathvik R. Kannan et al* » les chaînes latérales plus longues de l'Arginine par rapport à la leucine empêchent la liaison des anticorps avec le domaine **RBD** de la protéine **S**(109).

6.1.4 Mutation 478K, E484K/Q et N501Y

		480	490	500	
Chine/	471	E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T N			501
D614G/	471	E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T N			501
Alpha/	468	E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T Y			498
Eta/	468	E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T N			498
Béta/	470	E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y			500
GAMMA/	471	E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y			501
Delta/	469	E I Y Q A G S K P C N G V Q G F N C Y F P L Q S Y G F Q P T N			499
		T478K	E484K		N501Y
			E484Q		

Figure 17 : Les mutations T478K (Delta), E484K (Béta, Gamma, Eta), E484Q (Delta) et N501Y (Alpha, Béta, Gamma)

La figure 17 montre la présence : de la mutation **T478K** dans le variant Delta ; de la mutation **E484K** dans les variant Béta, Gamma, Eta ; de la mutation **E484Q** dans le variant Delta ; et de la mutation **N501Y** dans les variants Alpha, Béta, Gamma.

Selon les dernières données publiées par l’OMS ;

- La mutation **T478K** n’a été détecté que dans le variant Delta.
- La mutation **E484K** est la plus fréquente, elle est présente dans la majorité des variants ; Béta, Gamma, Eta, Iota, Mu, B.1.620.
- La mutation **E484Q** n’est présente que dans deux variants ; Delta (B.1.617.2, Kappa (B.1.617.1).
- La mutation **N501Y** est aussi très fréquentes, elle est détectée dans les variants ; Alpha, Béta, Gamma et Mu.

Selon « *Saathvik R. Kannan et al* », la mutation **T478K** (Variant Delta) entraîne une diminution de la neutralisation du virus par les anticorps monoclonaux, les sérums de convalescence et les sérums post –vaccination (109).

La mutation **T478K** perturbe la fixation des anticorps monoclonaux sur la protéine S qui est due à la différence dans la longueur des chaînes latérales entre la lysine (plus longue) et la thréonine (109).

La mutation **E484K** modifie la forme de l'extrémité de la protéine S entraînant une augmentation de l'affinité pour l'**ACE2** et permet aussi d'échapper partiellement aux anticorps principalement les anticorps poly-clonaux issus d'une infection ou vaccination antérieure (36,110).

La mutation **E484K** augmente sensiblement la quantité d'anticorps sériques nécessaires pour empêcher l'infection des cellules (111).

Selon « *Greaney et al* » la mutation **E484K** semble réduire significativement la neutralisation du virus par le sérum des convalescents(112).

Dans une étude menée par « *Nonaka et al.* », une femme précédemment infectée par un variant sans la mutation **E484Q** a été réinfecté par un variant présentant la mutation **E484K** (113).

Selon une autre étude publiée par « *Liu et al* », la mutation **E484K** entraîne la réduction de la neutralisation du virus par les anticorps monoclonaux (114).

La mutation **E484Q** est située au même endroit que **E484K** et par conséquent elle augmente l'affinité pour les récepteurs **ACE2** et peut aider le virus à échapper à certains types d'anticorps (107).

La mutation **N501Y** modifie légèrement l'extrémité de la protéine S ce qui augmente l'affinité de la protéine S pour les récepteurs **ACE2** améliorant ainsi l'attachement viral et l'entrée du virus dans les cellules hôtes par conséquent une augmentation de la transmission (36).

D'après une étude récente effectuée dans le centre de recherche de *Thomas J. Watson* à New-York, des chercheurs ont proposés une explication sur l'augmentation de la transmissibilité des variants portant la mutation **N501Y**. cette recherche suggère que la nature

hydrophobe de la protéine **Y501** (tyrosine) se lie plus étroitement aux fragments hydrophobes situés dans les récepteurs **ACE2** augmentant ainsi l'affinité entre la protéine **S** et les récepteurs **ACE2** et favorisant l'évasion aux anticorps, par contre la protéine **N501** est hydrophile (115).

Les mutations **K417N/T**, **L452R**, **E484K**, **E484Q** et **N501Y** sont situées au niveau du domaine **RBD** qui est constitué de 71 acides aminés situés entre la position 438 et 508, et est ainsi un point clé pour les mutations qui peuvent avoir une grande incidence clinique en termes de virulence, de transmissibilité et d'évasion immunitaire de l'hôte (116).

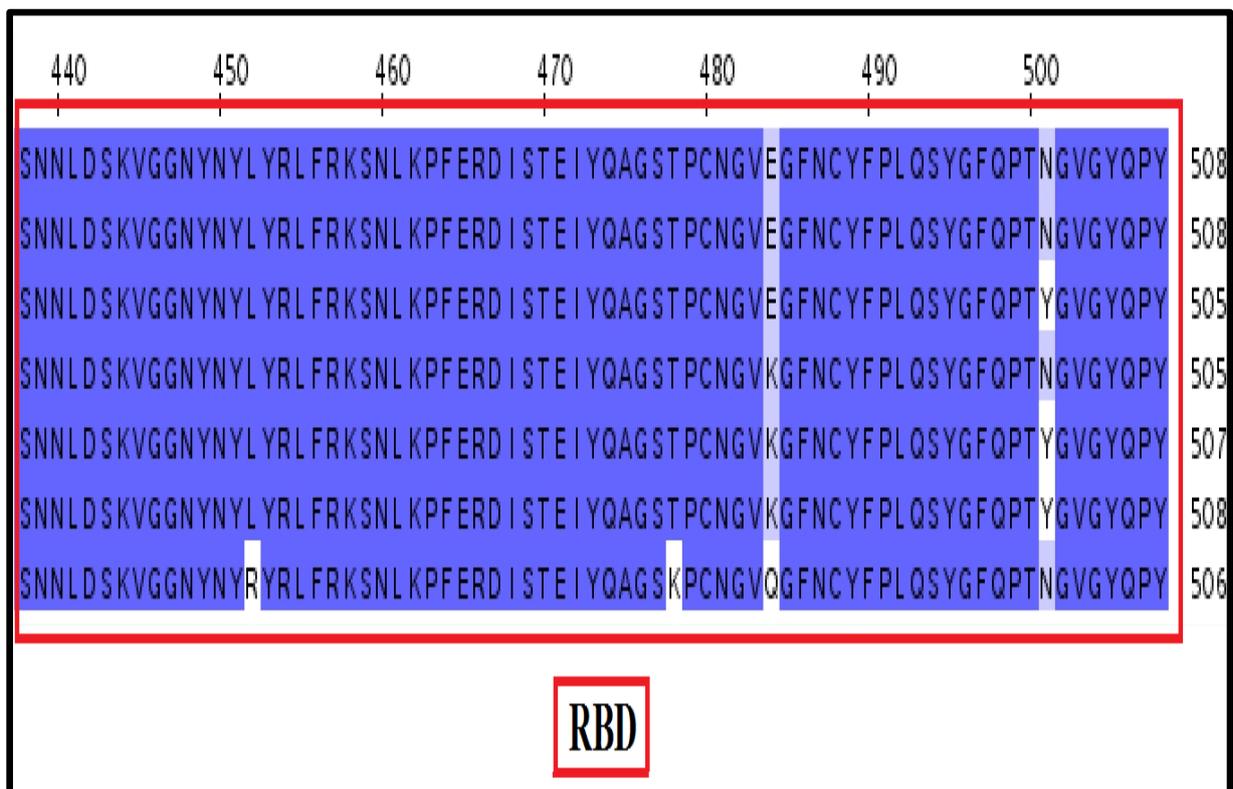


Figure 18 : Séquence protéique du domaine RBD.

La figure 18 représente le domaine **RBD** qui est situé entre la position 438 et 508 de la séquence protéique de la protéine **S**, nous pouvons aussi constater la présence des mutations ; **N501Y** dans la séquence du variant Alpha, **E484K** dans la séquence du variant Eta, **E484K** et **N501Y** dans la séquence du variant Béta et Gamma,

D'après la recherche publiée par « *Nelson et al* », les mutations **K417N**, **E484K** et **N501Y** augmente l'affinité de la protéine **S** pour les récepteurs **ACE2**, l'affinité est d'autant plus augmentée lorsque les deux mutations « **E484K** et **N501Y** » sont présentes dans la même variant. La mutation **K417N** semble stabiliser la mutation **N501Y** et ainsi stabiliser la liaison **protéine S – récepteur ACE2** (116).

6.1.5 La mutation D614G

		600	610	620	
<i>Chine/</i>	599	TPGTNTSNQVAVLYG D VNCTEVPVAIHADQL			629
<i>D614G/</i>	599	TPGTNTSNQVAVLYG G VNCTEVPVAIHADQL			629
<i>Alpha/</i>	596	TPGTNTSNQVAVLYG G VNCTEVPVAIHADQL			626
<i>Eta/</i>	596	TPGTNTSNQVAVLYG G VNCTEVPVAIHADQL			626
<i>Béta/</i>	598	TPGTNTSNQVAVLYG G VNCTEVPVAIHADQL			628
<i>GAMMA/</i>	599	TPGTNTSNQVAVLYG G VNCTEVPVAIHADQL			629
<i>Delta/</i>	597	TPGTNTSNQVAVLYG G VNCTEVPVAIHADQL			627

D614G

Figure 19 : Mutation D614G détectée dans les variants : Alpha, Béta Gamma, Delta, Eta et G614.

La figure 19 indique la présence de la mutation **D614G** dans les séquences des variants Alpha, Béta, Gamma, Delta, Eta et le variant G614 cependant la mutation est absente dans la séquence de la souche ancestrale.

La mutation **D614G** est située en dehors du site **RBD** et donc ne modifie pas l'affinité de la liaison avec les récepteurs **ACE2** ou la neutralisation du virus par les anticorps, cependant elle peut augmenter la densité de la protéine **S** en préservant son intégrité structurelle et évite la perte de la sous unité S1 entraînant en conséquence une augmentation de la transmissibilité du virus ce qui explique le fait qu'elle soit la mutation la plus répandue dans le monde (actuellement détectée dans presque tous les variants en circulation) et qu'elle est remplacée la souche ancestrale (117).

Récemment, des chercheurs ont identifié des mutations potentiellement dangereuses dans deux des domaines fonctionnels de la protéine virale **S** ;

- Le domaine **RBD** ; R408I (Arginine à Isoleucine), L455Y (Leucine à Tyrosine), F486L (phénylalanine à leucine), Q493N, Q498Y et N501Y (118).

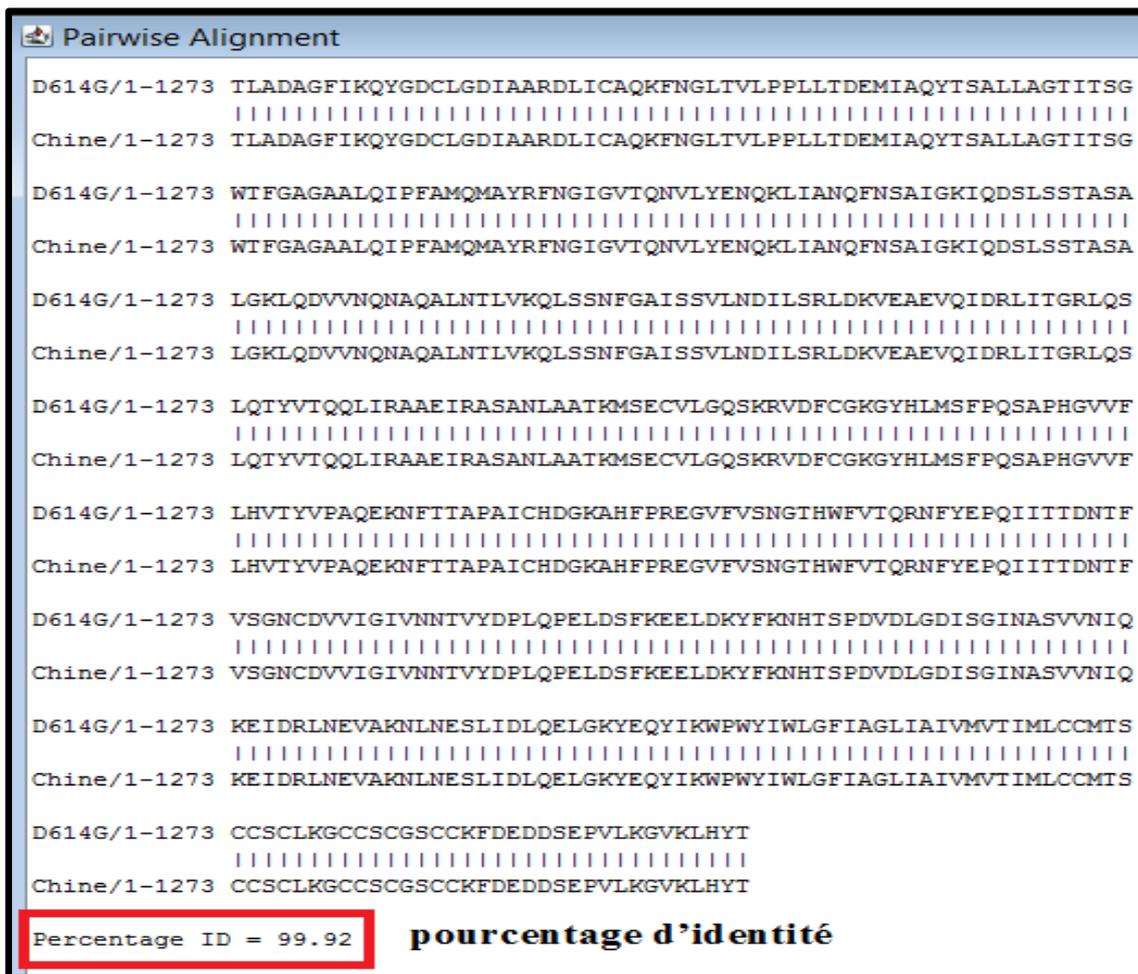
- Le domaine **HR1** (Heptad-repeat) ; A930V (Alanine à valine) , D936Y.(Acide aspartique à tyrosine) (118).

6.2 Pourcentage d'identité

L'alignement multiple nous permet aussi de mesurer le pourcentage d'identité des souches par rapport à la souche initiale ainsi qu'entre les souches elles-mêmes, ça nous permet d'évaluer le degré l'évolution génétique des variants du SARS-COV-2 (119).

Le calcul du pourcentage d'identité est effectué par le logiciel « JALVIEW », les étapes sont assez simples ;

- Sélectionner deux séquences à aligner puis appuyer sur l'onglet « **Calculate** ».
- Choisir l'option « **Pairwise alignment** » et une nouvelle console s'ouvre qui affiche l'alignement ainsi que le pourcentage d'identité (Figure 20).



```

Pairwise Alignment
D614G/1-1273 TLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSG
Chine/1-1273 TLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSG

D614G/1-1273 WTFGAGAALQIPFAMQMAYRFNGIGVTVQNVLYENQKLIANQFNSAIGKIQDSLSSTASA
Chine/1-1273 WTFGAGAALQIPFAMQMAYRFNGIGVTVQNVLYENQKLIANQFNSAIGKIQDSLSSTASA

D614G/1-1273 LGKLDVNVQNAQALNTLVKQLSSNFGAIISSVLNDILSRLDKVEAEVQIDRLITGRLQS
Chine/1-1273 LGKLDVNVQNAQALNTLVKQLSSNFGAIISSVLNDILSRLDKVEAEVQIDRLITGRLQS

D614G/1-1273 LQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
Chine/1-1273 LQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF

D614G/1-1273 LHVTYVPAQEKNFTTAPAICHGKAHFPPREGVFSNGTHWFVTQRNFYEPQIITTDNTE
Chine/1-1273 LHVTYVPAQEKNFTTAPAICHGKAHFPPREGVFSNGTHWFVTQRNFYEPQIITTDNTE

D614G/1-1273 VSGNCDVVIGIVNNTIVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQ
Chine/1-1273 VSGNCDVVIGIVNNTIVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQ

D614G/1-1273 KEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTS
Chine/1-1273 KEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTS

D614G/1-1273 CCSCLKGCCSCGSCCKFDEDDSEPVKGVKLYHT
Chine/1-1273 CCSCLKGCCSCGSCCKFDEDDSEPVKGVKLYHT

Percentage ID = 99.92
pourcentage d'identité

```

Figure 20 : Alignement de séquence deux à deux et pourcentage d'identité.

La figure 20 montre l'alignement deux à deux des séquences sélectionnées ainsi que le pourcentage d'identité entre ces deux séquences.

En plus de l'alignement des séquences de la protéine S, nous avons décidé aussi de réaliser l'alignement des séquences des domaines où les mutations sont les plus fréquentes afin d'obtenir une image globale sur l'évolution génétique du SARS-COV-2 par des mutations de plus en plus fréquentes.

Le tableau ci-dessous présente le pourcentage d'identité entre le génome des variants par rapport à celui de la souche ancestrale.

Tableau 3 : pourcentage d'identité des différents domaines du SARS-COV-2 des variants par rapport à la souche ancestrale.

Pourcentage d'identité							
Domaines Souches	Génome complet	Protéine S	RBD	ORF1ab	ORF1a	Protéine N	ORF3a
Wuhan, Chine (ancestrale)	100%	100%	100%	100%	100%	100%	100%
G614	99.97%	99.92%	100%	99.97%	99.98%	100%	99.64%
Variant Alpha	99.89%	99.14%	98.59%	99.93%	99.91%	98.81%	99,64%
Variant Béta	99.89%	99.06%	97.18%	99.86%	99.82%	99.76%	99.27%
Variant Gamma	99.87%	99.98%	97.18%	99.90%	99.89%	99.28%	99.64%
Variant Delta	99.89%	99,21%	95.77%	99.87%	99.89%	98.81%	99.64%
Variant Eta	99.91%	99.21%	98.59%	99.94%	99.95%	99.04%	99.64%

La couleur verte représente la souche ancestrale.

La couleur bleue indique les pourcentages les plus élevés.

✚ La couleur rouge indique les pourcentages les moins élevés.

Le variant G614 qui est aussi le premier variant du SARS-COV-2 possède le pourcentage d'identité le plus proche de celui de la souche ancestrale.

Le variant Béta est celui qui présente des pourcentages d'identités les moins élevés suivi du variant Delta.

Le domaine ORF3a présente un pourcentage d'identité très stable entre les variants ce qui indique que c'est un domaine conservé, ce résultat est en accord avec les résultats obtenus dans la recherche menés par *Martina Bianchi et al.* (120)

Le domaine **RBD** est celui qui possède les pourcentages d'identité les plus disparates entre les variants et la souche ancestrale par rapport aux autres domaines du génome viral, indiquant ainsi qu'il est le domaine qui a subi le plus de variations génétiques. Ces résultats vont dans le sens de l'idée formulée dans ce travail qui est que les mutations tendent à s'accumuler dans le domaine **RBD** ce qui peut éventuellement expliquer la raison pour laquelle les variants sont plus transmissibles et peuvent échapper au système immunitaire.

7 Conclusion

Tous les virus y compris le SARS-CoV-2 mutent avec le temps, la plupart des mutations n'ont que peu ou pas d'incidence sur les propriétés du virus. Cependant certaines mutations sont plus inquiétantes et constituent un défi pour la prévention et le traitement du SARS-COV-2. La plupart des vaccins contre le SARS-COV-2 approuvés ou qui sont encore en cours d'essai clinique sont conçus sur la base de la protéine **S**. cependant la fréquence de mutation de la protéine **S** est très élevée car elle est en première ligne contre le système immunitaire ce qui l'oblige à s'adapter et à évoluer par des mutations lui conférant des capacités de neutralisation et d'évasion immunitaire qui rendent la tâche de confection des vaccins et traitements plus difficile.

A travers ce travail nous avons pu constater que le SARS-COV-2 évolue par des mutations au niveau de son génome, ces mutations peuvent se produire sur la totalité du génome viral tout comme elles peuvent se produire dans certaines régions bien spécifiques du génome. La majorité des mutations détectées s'accumulent au niveau des domaines : **ORF1, ORF3, ORF8**, protéine **N**. **PNS2, PNS3, PNS16 (RdRp)**, et plus particulièrement dans la protéine **S**. Par contre certains domaines n'ont à ce jour présenté aucune ou peu de mutations tels que ; PNS7, PNS9, PNS10, PNS11, protéine **E**, ORF6, ce qui suggère que ces protéines ont des fonctions conservées qui sont essentielles à la propagation du virus.

La protéine **S** compte 60 mutations réparties sur la totalité de son code génétique, parmi ces mutations celles situées dans le domaine **RBD** « **K417N/T, L452R, K478T, E484K, E484Q** et **N501Y** » sont celles qui ont eu le plus d'impact sur la virulence du virus, sur sa transmissibilité ainsi que sur sa capacité à échapper au système immunitaire, par conséquent les variants qui abritent ces mutations sont les plus dominants et les plus contagieux au monde tout particulièrement le variant **Alpha (E484K et N501Y)** et **Delta (L452R et K478T)** qui sont actuellement dominants en Algérie.

L'analyse systématique *in silico* de l'évolution du génome du SARS-CoV-2 est un outil puissant pour obtenir des éléments de compréhension de la biologie moléculaire et de la pathogénie virale dont l'un des principaux buts est la conception d'expériences ou de stratégies thérapeutiques et vaccinales spécifiques.

Bien que l'idée de départ était d'étudier les séquences protéiques des variants détectés en Algérie mais devant la non disponibilité des séquences on s'est rabattu sur les séquences accessibles en ligne dans les bases de données protéiques et nucléiques. Par conséquent il serait intéressant que ce travail soit repris en exploitant les données locales des échantillons des variants séquencés en Algérie et de faire une analyse comparative avec les séquences publiées dans les autres pays et ainsi étudier l'impact de la géographie et du climat sur l'évolution du virus et l'adaptation du virus au système immunitaire et vice versa.

Références bibliographiques

1. Fagbule OF. 2019 NOVEL CORONAVIRUS. *Ann Ib Postgrad Med.* 2019 Dec;17(2):108–10.
2. Haque SM, Ashwaq O, Sarief A, Azad John Mohamed AK. A comprehensive review about SARS-CoV-2. *Future Virol.* 2020 Sep 1;15(9):625–48.
3. Konings F, Perkins MD, Kuhn JH, Pallen MJ, Alm EJ, Archer BN, et al. SARS-CoV-2 Variants of Interest and Concern naming scheme conducive for global discourse. *Nat Microbiol.* 2021 Jul;6(7):821–3.
4. Gómez CE, Perdiguero B, Esteban M. Emerging SARS-CoV-2 Variants and Impact in Global Vaccination Programs against SARS-CoV-2/COVID-19. *Vaccines.* 2021 Mar 11;9(3):243.
5. Liu Z, Xiao X, Wei X, Li J, Yang J, Tan H, et al. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J Med Virol.* 2020 Mar 11.
6. Jin Y, Yang H, Ji W, Wu W, Chen S, Zhang W, et al. Virology, Epidemiology, Pathogenesis, and Control of COVID-19. *Viruses.* 2020 Mar 27 ;12(4).
7. Vabret A, Mischczak F. Changement de tropisme des coronavirus. *Rev Francoph Lab.* 2010 Jun;2010(423):63–8.
8. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol.* 2016 Jun;24(6):490–502.
9. Sood S, Aggarwal V, Aggarwal D, Upadhyay SK, Sak K, Tuli HS, et al. COVID-19 Pandemic: from Molecular Biology, Pathogenesis, Detection, and Treatment to Global Societal Impact. *Curr Pharmacol Rep.* 2020 Jul 27;1–16.
10. Vabret A, Dina J, Brison E, Brouard J, Freymuth F. Coronavirus humains (HCoV). *Pathol Biol (Paris).* 2009 Mar;57(2):149–60.
11. Juckel D, Dubuisson J, Belouzard S. Les coronavirus, ennemis incertains. *médecine/sciences.* 2020 Jun 1;36(6–7):633–41.
12. Wang M-Y, Zhao R, Gao L-J, Gao X-F, Wang D-P, Cao J-M. SARS-CoV-2: Structure,

Biology, and Structure-Based Therapeutics Development. *Front Cell Infect Microbiol.* 2020

Nov 25

13. Helmy YA, Fawzy M, Elasad A, Sobieh A, Kenney SP, Shehata AA. The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. *J Clin Med.* 2020 Apr 24;9(4):1225.
14. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. *Virology.* 2019 May 27;16(1):69.
15. Masters PS. The Molecular Biology of Coronaviruses. *Adv Virus Res.* 2006;66:193–292.
16. Machhi J, Herskovitz J, Senan AM, Dutta D, Nath B, Oleynikov MD, et al. The Natural History, Pathobiology, and Clinical Manifestations of SARS-CoV-2 Infections. *J Neuroimmune Pharmacol.* 2020 Jul 21;1–28.
17. Enjuanes Luis, Coronaviridae - coronavirus:
<http://virologie.free.fr/documents/virologie/34-Coronaviridae/coronaviridae.htm>
18. NIAID. Novel Coronavirus SARS-CoV-2. 2020 <https://www.niaid.nih.gov/news-events/novel-coronavirus-sarscov2-images>
19. Cascella M, Rajnik M, Aleem A, Dulebohn SC, Di Napoli R. Features, Evaluation, and Treatment of Coronavirus (COVID-19). In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2021 [cited 2021 Sep 16].
20. Sophie T. Aspects immunologiques et virologiques de l'infection par le SARS-CoV-2. 2020 ;134.
21. Zhou H, Chen X, Hu T, Li J, Song H, Liu Y, et al. A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the Spike Protein. *Curr Biol.* 2020 Jun 8;30(11):2196-2203.e3.
22. Maleki M. chapter 40 - COVID-19 Infection: A Novel Fatal Pandemic of the World in 2020. :5.
23. Zayratyants OV, Samsonova MV, Mikhaleva LM, Chernyaev AL, Mishnev OD, Krupnov NM, et al. Introduction. *Pathology of COVID-19: Atlas*. Research Institute for Healthcare Organization and Medical Management, Moscow Healthcare Department; 2020
24. Wang R, Pan M, Zhang X, Han M, Fan X, Zhao F, et al. Epidemiological and clinical features

- of 125 Hospitalized Patients with COVID-19 in Fuyang, Anhui, China. *Int J Infect Dis.* 2020 Jun; 95:421–8.
25. Hardy ÉJL, Flori P. Spécificités épidémiologiques de la COVID-19 en Afrique : préoccupation de santé publique actuelle ou future ? *Ann Pharm Fr.* 2021 Mar;79(2):216–26.
26. Rosenthal PJ, Breman JG, Djimde AA, John CC, Kanya MR, Leke RGF, et al. COVID-19: Shining the Light on Africa. *Am J Trop Med Hyg.* 2020 Jun 3;102(6):1145–8.
27. Tu H, Tu S, Gao S, Shao A, Sheng J. Current epidemiological and clinical features of COVID-19; a global perspective from China. *J Infect.* 2020 Jul;81(1):1–9.
28. Wu D, Wu T, Liu Q, Yang Z. The SARS-CoV-2 outbreak: What we know. *Int J Infect Dis.* 2020 May;94:44–8.
29. Ye Z-W, Yuan S, Yuen K-S, Fung S-Y, Chan C-P, Jin D-Y. Zoonotic origins of human coronaviruses. *Int J Biol Sci.* 2020 Mar 15;16(10):1686–97.
30. Segondy M. Les Coronavirus humains. *Rev Francoph Lab.* 2020 Nov;2020(526):32–9.
31. Corum J, Zimmer C. Coronavirus Variants and Mutations. *The New York Times*; <https://www.nytimes.com/interactive/2021/health/coronavirus-variant-tracker.html>
32. Salzberger B, Buder F, Lampl B, Ehrenstein B, Hitzenbichler F, Holzmann T, et al. Epidemiology of SARS-CoV-2. *Infection.* 2021;49(2):233–9.
33. Konings F, Perkins MD, Kuhn JH, Pallen MJ, Alm EJ, Archer BN, et al. SARS-CoV-2 Variants of Interest and Concern naming scheme conducive for global discourse. *Nat Microbiol.* 2021 Jul;6(7):821–3.
34. Guo S, Liu K, Zheng J. The Genetic Variant of SARS-CoV-2: would It Matter for Controlling the Devastating Pandemic? *Int J Biol Sci.* 2021 Apr 10;17(6):1476–85.
35. Chen Y, Liu Q, Guo D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol.* 2020 Apr;92(4):418–23.
36. Larsen CS, Paludan SR. Corona's new coat: SARS-CoV-2 in Danish minks and implications for travel medicine. *Travel Med Infect Dis.* 2020;38:101922.
37. Tang JW, Tambyah PA, Hui DS. Emergence of a new SARS-CoV-2 variant in the UK. *J*

Infect. 2021 Apr;82(4):e27–8.

38. Meyerowitz EA, Richterman A, Gandhi RT, Sax PE. Transmission of SARS-CoV-2: A Review of Viral, Host, and Environmental Factors. *Ann Intern Med.* 2020 Sep 17

39. Fiolet Thibault, Coronavirus 2019-nCoV, taux de reproduction de base et létalité : comment évaluer les risques avec l'épidémiologie ? . Quoi dans mon assiette. 2020: <https://quoidansmonassiette.fr/coronavirus-2019-risques-deces-ncov-mortalite-taux-de-reproduction-de-base-et-letalite-evaluer-les-risques/>

40. Berry Emi. What we know about the SARS-CoV-2 Delta variant. UNSW Newsroom.2021: <https://newsroom.unsw.edu.au/news/health/what-we-know-about-sars-cov-2-delta-variant>

41. COVID Live Update - Worldometer. https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1

42. Segondy M. Les Coronavirus humains. *Rev Francoph Lab.* 2020 Nov;2020(526):32–9.

43. Mukherjee S, Pahan K. Is COVID-19 Gender-sensitive? *J Neuroimmune Pharmacol.* 2021 Mar 1;16(1):38–47.

44. Maleki Dana P, Sadoughi F, Hallajzadeh J, Asemi Z, Mansournia MA, Yousefi B, et al. An Insight into the Sex Differences in COVID-19 Patients: What are the Possible Causes? *Prehospital Disaster Med.* :1–4.

45. Chakravarty D, Nair SS, Hammouda N, Ratnani P, Gharib Y, Wagaskar V, et al. Sex differences in SARS-CoV-2 infection rates and the potential link to prostate cancer. *Commun Biol.* 2020 Jul 8;3(1):1–12.

46. Poletti P, Tirani M, Cereda D, Trentini F, Guzzetta G, Marziano V, et al. Age-specific SARS-CoV-2 infection fatality ratio and associated risk factors, Italy, February to April 2020. *Eurosurveillance.* 2020 Aug 6;25(31).

47. Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older people? *Aging.* 2020 May 29 ;12(10):9959–81.

48. Nikolich-Zugich J, Knox KS, Rios CT, Natt B, Bhattacharya D, Fain MJ. SARS-CoV-2 and COVID-19 in older adults: what we may expect regarding pathogenesis, immune responses, and outcomes. *GeroScience.* 2020 Apr 10;42(2):505–14.

49. Hartenian E, Nandakumar D, Lari A, Ly M, Tucker JM, Glaunsinger BA. The molecular virology of coronaviruses. *J Biol Chem.* 2020 Sep 11;295(37):12910–34.
50. Rabaan AA, Al-Ahmed SH, Haque S, Sah R, Tiwari R, Malik YS, et al. SARS-CoV-2, SARS-CoV, and MERS-CoV: a comparative overview. :11.
51. Sanyal S. How SARS-CoV-2 (COVID-19) spreads within infected hosts — what we know so far. *Emerg Top Life Sci.* 2020 Dec 11;4(4):383–90.
52. Fung S-Y, Yuen K-S, Ye Z-W, Chan C-P, Jin D-Y. A tug-of-war between severe acute respiratory syndrome coronavirus 2 and host antiviral defence: lessons from other pathogenic viruses. *Emerg Microbes Infect.* 2020 Mar 14;9(1):558–70.
53. Mohamadian M, Chiti H, Shoghli A, Biglari S, Parsamanesh N, Esmailzadeh A. COVID-19: Virology, biology and novel laboratory diagnosis. *J Gene Med.* 2021 Feb;23(2).
54. Rahimi A, Mirzazadeh A, Tavakolpour S. Genetics and genomics of SARS-CoV-2: A review of the literature with the special focus on genetic diversity and SARS-CoV-2 genome detection. *Genomics.* 2021 Jan;113(1):1221–32.
55. Ellis, P., Somogyvári, F., Virok, D.P. *et al.* Decoding Covid-19 with the SARS-CoV-2 Genome. *Curr Genet Med Rep* **9**, 1–12 (2021).
56. Saxena SK, Kumar S, Baxi P, Srivastava N, Puri B, Ratho RK. Chasing COVID-19 through SARS-CoV-2 spike glycoprotein. *Virus Disease.* 2020 Dec;31(4):399–407.
57. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol.* 2020 Oct 28;1–16.
58. SARS-CoV-2 - Antigènes (Protéines et peptides) pour la recherche et : <https://www.clinisciences.com/achat/cat-sars-cov-2-antigenes-proteines-5102.html>
59. Zhou W, Wang W. Fast-spreading SARS-CoV-2 variants: challenges to and new design strategies of COVID-19 vaccines. *Signal Transduct Target Ther.* 2021 Jun 9;6(1):1–6.
60. Bamford C. New coronavirus variant: what is the spike protein and why are mutations on it important? . *The Conversation* : <http://theconversation.com/new-coronavirus-variant-what-is-the-spike-protein-and-why-are-mutations-on-it-important-152463>
61. Thomas S. The Structure of the Membrane Protein of SARS-CoV-2 Resembles the Sugar Transporter SemiSWEET. *Pathog Immun.* 2020 Oct 19;5(1):342–63.
62. SATIJA N, LAL SK. The Molecular Biology of SARS Coronavirus. *Ann N Y Acad Sci.* 2007 Apr;1102(1):26–38.
63. Fehr AR, Perlman S. Coronaviruses: An Overview of Their Replication and

- Pathogenesis. Coronaviruses. 2015 Feb 12;1282:1–23.
64. SARS-CoV-2 Research Resources for Non-Structural and Accessory Proteins. Novus Biologicals.: <https://www.novusbio.com/support/sars-cov-research-resources/nonstructural-proteins>
65. Harrison AG, Lin T, Wang P. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. Trends Immunol. 2020 Dec;41(12):1100–15.
66. Trougakos IP, Stamatelopoulos K, Terpos E, Tsitsilonis OE, Aivalioti E, Paraskevis D, et al. Insights to SARS-CoV-2 life cycle, pathophysiology, and rationalized treatments that target COVID-19 clinical complications. J Biomed Sci. 2021 Jan 12;28.
67. Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan ES, Izard T, et al. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. bioRxiv. 2020 Jun 12
68. Zhang Y, Xi H, Juhas M. Biosensing Detection of the SARS-CoV-2 D614G Mutation. Trends Genet. 2021 Apr 1;37(4):299–302.
69. CDC. Coronavirus Disease 2019 (COVID-19). Centers for Disease Control and Prevention. 2020: <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html>
70. Features, Evaluation, and Treatment of Coronavirus (COVID-19) - StatPearls - NCBI Bookshelf.
71. Sieńko J, Kotowski M, Bogacz A, Lechowicz K, Drożdżal S, Rosik J, et al. COVID-19: The Influence of ACE Genotype and ACE-I and ARBs on the Course of SARS-CoV-2 Infection in Elderly Patients. Clin Interv Aging. 2020 Jul 21;15:1231–40.
72. Gemmati D, Bramanti B, Serino ML, Secchiero P, Zauli G, Tisato V. COVID-19 and Individual Genetic Susceptibility/Receptivity: Role of ACE1/ACE2 Genes, Immunity, Inflammation and Coagulation. Might the Double X-Chromosome in Females Be Protective against SARS-CoV-2 Compared to the Single X-Chromosome in Males? Int JMol Sci. 2020 May 14;21(10):3474.
73. Zipeto D, Palmeira J da F, Argañaraz GA, Argañaraz ER. ACE2/ADAM17/TMPRSS2 Interplay May Be the Main Risk Factor for COVID-19. Front Immunol. 2020 Oct 7;11:576745.
74. Iran reports 116 deaths and 2,139 new coronavirus cases within 24 hours - Europe & World News: <https://www.europeworldnews.com/iran-reports-116-deaths-and-2139-new-coronavirus-cases-within-24-hours/>
75. Mécanisme de RNAi. News-Medical.net. 2018: [https://www.news-medical.net/life-sciences/RNAi-mechanism-\(French\).aspx](https://www.news-medical.net/life-sciences/RNAi-mechanism-(French).aspx)
76. Bonny V, Maillard A, Mousseaux C, Plaçais L, Richier Q. COVID-19 :

- physiopathologie d'une maladie à plusieurs visages. *Rev Med Interne*. 2020 Jun;41(6):375–89.
77. Deng L, Zeng Q, Wang M, Cheng A, Jia R, Chen S, et al. Suppression of NF- κ B Activity: A Viral Immune Evasion Mechanism. *Viruses*. 2018 Aug 4;10(8)
78. Tregoning JS, Schwarze J. Respiratory Viral Infections in Infants: Causes, Clinical Symptoms, Virology, and Immunology. *Clin Microbiol Rev*. 2010 Jan;23(1):74–98.
79. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun*. 2020 May ;109 :102433.
80. Hantz S. Diagnostic biologique de l'infection à Sars-CoV-2 : stratégies et interprétation des résultats. *Rev Francoph Lab*. 2020 Nov ;2020(526):48–56.
81. Deluzarche C. La Covid-19 peut encore tuer six mois après l'infection. *Futura*.: <https://www.futura-sciences.com/sante/actualites/coronavirus-covid-19-peut-encore-tuer-six-mois-apres-infection-85150/>
82. Gala J, Nyabi O, Durant J.F, Chibani N, Bentahir M, Méthodes diagnostiques du COVID-19 | *Louvain Médical*. <https://www.louvainmedical.be/fr/article/methodes-diagnostiques-du-covid-19>
83. Desvaux É, Faucher J-F. Covid-19 : aspects cliniques et principaux éléments de prise en charge. *Rev Francoph Lab*. 2020 Nov;2020(526):40–7.
84. Kumar R, Nagpal S, Kaushik S, Mendiratta S. COVID-19 diagnostic approaches: different roads to the same destination. *VirusDisease*. 2020 Jun 1;31(2):97–105.
85. Udugama B, Kadhiresan P, Kozłowski HN, Malekjahani A, Osborne M, Li VYC, et al. Diagnosing COVID-19: The Disease and Tools for Detection. *ACS Nano*. 2020 Mar 30
86. Hu B, Guo H, Zhou P, Shi Z-L. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol*. 2020 Oct 6;1–14.
87. Matusik É, Ayadi M, Picard N. Covid-19, prise en charge, pistes thérapeutiques et vaccinales. *Actual Pharm*. 2020 Oct;59(599):27–33.
88. Ng TSB, Leblanc K, Yeung DF, Tsang TSM. Médicaments utilisés durant la COVID-19. *Can Fam Physician*. 2021 Mar;67(3):e69–78.
89. Canedo-Marroquín G, Saavedra F, Andrade CA, Berrios RV, Rodríguez-Guilarte L, Opazo MC, et al. SARS-CoV-2: Immune Response Elicited by Infection and Development of Vaccines and Treatments. *Front Immunol*. 2020 Dec 11 11
90. Axfors C, Schmitt AM, Janiaud P, van't Hooft J, Abd-Elsalam S, Abdo EF, et al. Mortality outcomes with hydroxychloroquine and chloroquine in COVID-19 from an international

- collaborative meta-analysis of randomized trials. *Nat Commun.* 2021 Apr 15; 12:2349.
91. Zuily S, Dufrost V, Wahl D. Thrombose artérielle et veineuse au cours du COVID-19. *Arch Mal Coeur Vaiss Prat.* 2020 Oct;2020(291):16–9.
92. Peiffer-Smadja N, Rozencwajg S, Kherabi Y, Yazdanpanah Y, Montravers P. Vaccins COVID-19 : une course contre la montre. *Anesth Réanimation.* 2021 May;7(3):199–202.
93. Mulder N, Schwartz R, Brazas MD, Brooksbank C, Gaeta B, Morgan SL, et al. The development and application of bioinformatics core competencies to improve bioinformatics training and education. *PLoS Comput Biol.* 2018 Feb 1;14(2):e1005772.
94. Moore JH. Bioinformatics. *J Cell Physiol.* 2007;213(2):365–9.
95. Madeira F, Park Y mi, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 2019 Jul 2 ;47(W1):W636–41.
96. GenBank Overview: <https://www.ncbi.nlm.nih.gov/genbank/>
97. Mashima J, Kodama Y, Kosuge T, Fujisawa T, Katayama T, Nagasaki H, et al. DNA data bank of Japan (DDBJ) progress report. *Nucleic Acids Res.* 2016 Jan 4;44(Database issue): D51–7.
98. Barker WC, Garavelli JS, Haft DH, Hunt LT, Marzec CR, Orcutt BC, et al. The PIR-International Protein Sequence Database. *Nucleic Acids Res.* 1998 Jan 1;26(1):27–32.
99. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 2000 Jan 1;28(1):45–8.
100. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, et al. The UCSC Genome Browser Database. *Nucleic Acids Res.* 2003 Jan 1;31(1):51–4.
101. Mezhoud K. Alignement de séquences Principes et méthodes. :69.
102. Ray M, Sable MN, Sarkar S, Hallur V. Essential interpretations of bioinformatics in COVID-19 pandemic. *Meta Gene.* 2021 Feb ;27:100844.
103. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol.* 1970 Mar 28 ;48(3):443–53.
104. Higgins DG, Sharp PM. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene.* 1988 Dec 15;73(1):237–44.
105. Kemp SA, Meng B, Ferreira IA, Datir R, Harvey WT, Papa G, et al. Recurrent emergence and transmission of a SARS-CoV-2 spike deletion H69/V70. 2021 Mar p. 2020.12.14.422555.
106. Takano K, Watanabe Y, Hariu M, Seki M. Detection of Representative Mutant Strains and a Case of Prolonged Infection by SARS-CoV-2 with Spike 69/70 Deletion in Japan. *Infect Drug Resist.* 2021 Jul 6; 14:2579–81.

107. Boehm E, Kronig I, Neher RA, Eckerle I, Vetter P, Kaiser L. Novel SARS-CoV-2 variants: the pandemics within the pandemic. *Clin Microbiol Infect.* 2021 May 17
108. Fratev F. The N501Y and K417N mutations in the spike protein of SARS-CoV-2 alter the interactions with both hACE2 and human derived antibody: A Free energy of perturbation study. 2020 Dec
109. Kannan SR, Spratt AN, Cohen AR, Naqvi SH, Chand HS, Quinn TP, et al. Evolutionary analysis of the Delta and Delta Plus variants of the SARS-CoV-2 viruses. *J Autoimmun.* 2021 Nov; 124:102715.
110. Jangra S, Ye C, Rathnasinghe R, Stadlbauer D, Krammer F, Simon V, et al. The E484K mutation in the SARS-CoV-2 spike protein reduces but does not abolish neutralizing activity of human convalescent and post-vaccination sera. *medRxiv.* 2021 Jan 29;2021.01.26.21250543.
111. Wise J. Covid-19: The E484K mutation and the risks it poses. *BMJ.* 2021 Feb 5;372: n359.
112. Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, Loes AN, et al. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe.* 2021 Jan 13;29(1):44-57.e9.
113. Nonaka CKV, Franco MM, Gräf T, Mendes AVA, Aguiar RS de, Giovanetti M, et al. Genomic Evidence of a Sars-Cov-2 Reinfection Case With E484K Spike Mutation in Brazil. 2021 Jan 6
114. Xie X, Liu Y, Liu J, Zhang X, Zou J, Fontes-Garfias CR, et al. Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. *Nat Med.* 2021 Apr;27(4):620–1.
115. Luan B, Wang H, Huynh T. Molecular Mechanism of the N501Y Mutation for Enhanced Binding between SARS-CoV-2's Spike Protein and Human ACE2 Receptor. 2021 Jan p. 2021.01.04.425316.
116. Nelson G, Buzko O, Spilman P, Niazi K, Rabizadeh S, Soon-Shiong P. Molecular dynamic simulation reveals E484K mutation enhances spike RBD-ACE2 affinity and the combination of E484K, K417N and N501Y mutations (501Y.V2 variant) induces conformational change greater than N501Y mutant alone, potentially resulting in an escape mutant. 2021 Jan

117. Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan ES, IZard T, et al. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. *bioRxiv*. 2020 Jun 12;2020.06.12.148726.
118. Ahamad S, Kanipakam H, Gupta D. Insights into the structural and dynamical changes of spike glycoprotein mutations associated with SARS-CoV-2 host receptor binding. *J Biomol Struct Dyn*. :1–13.
119. Pirovano W, Heringa J. Multiple Sequence Alignment. In: Keith JM, editor. *Bioinformatics: Data, Sequence Analysis and Evolution*. Totowa, NJ: Humana Press; 2008. p. 143–61. (Methods in Molecular Biology™).
120. Bianchi M, Borsetti A, Ciccozzi M, Pascarella S. SARS-Cov-2 ORF3a: Mutability and function. *Int J Biol Macromol*. 2021 Feb 15; 170:820–6.

Soutenu le: 09/2021	présenté par : KSIR NAZIM
<i>Etude in silico des mutations de la protéine S « spike protein » détectées dans les principaux variants du SARS-CoV-2 aperçus en Algérie.</i>	
<p>Résumé</p> <p>L'émergence de la maladie du syndrome respiratoire aigu sévère à coronavirus 2 (SARS-COV-2) en Chine fin 2019 a provoqué une vaste épidémie mondiale et constitue un problème majeur pour la santé publique. Malgré la lenteur de l'évolution du SARS-COV-2 par rapport aux autres virus à ARN, sa transmission massive et rapide pendant la pandémie mondiale lui a permis d'évoluer par l'acquisition d'une diversité génétique importante qui a conduit à l'émergence de nombreux variants, parmi lesquelles certains possèdent des mutations inquiétantes qui semblent avoir un impact sur la transmissibilité du virus, les manifestations cliniques, la mortalité ainsi que sur l'évasion immunitaire qui constitue le principale frein au développement d'un vaccin efficace contre les différentes souches du SARS-COV-2. Actuellement le GISAID (Global Initiative on Sharing Avian Influenza Data) a répertorié plus de 250 mutations au niveau du génome du SARS-COV-2 dont 60 mutations au niveau de la protéine S, cette fréquence élevée de mutations au niveau de la protéine S reflète son importance d'où l'intérêt de l'étudier.</p> <p>Le présent travail a pour objectif d'étudier la protéine S et les mutations détectées dans sa séquence génomique, pour cela nous allons utiliser les séquences protéiques de la protéine S des différentes souches du SARS-COV-2 présentes dans les bases de données génomiques et les traiter par un algorithme d'alignement multiple pour les aligner et les comparer entre elles et par rapport à la souche ancestrale. Les résultats sont assez explicites en eux même et sont conformes aux données scientifiques des recherches déjà effectuées à savoir que la fréquence de mutation de la protéine S est plus élevée par rapport aux autres domaines du génome du SARS-COV-2 et que les mutations qui s'y produisent tout particulièrement au niveau du domaine RBD sont les plus dangereuses et qui ont eu le plus d'impacts sur la transmissibilité, l'évasion immunitaire et sur l'efficacité des vaccins.</p>	
<p>Mots clés :</p> <p>Coronavirus, COVID19, SARS-COV-2, Variant, ACE2, Protéine S, RBD, alignement séquence.</p>	
<p>Département de biologie animale</p> <p>Faculté des sciences de la nature et de la vie</p> <p>Université des frères constantine 1</p>	
<p><i>Directrice de thèse : OTMANI KHAWLA</i> <i>Examinatrice : GUENDOUZE ASSIA</i></p> <p style="text-align: center;"><i>Présidente du jury : SEDRATI KHADIDJA</i></p>	